



ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

## STOCK MARKET ΜΕ ΧΡΗΣΗ LLM

ΧΡΗΣΤΟΣ ΜΗΤΣΟΥΛΗΣ  
ΙΩΑΝΝΗΣ ΠΕΛΕΚΑΝΟΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΥΠΕΥΘΥΝΟΣ

ΚΩΝΣΤΑΝΤΙΝΟΣ ΚΟΛΟΜΒΑΤΣΟΣ  
ΑΝΑΠΛΗΡΩΤΗΣ ΚΑΘΗΓΗΤΗΣ

Λαμία 2025-2026





ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

# STOCK MARKET ΜΕ ΧΡΗΣΗ LLM

ΧΡΗΣΤΟΣ ΜΗΤΣΟΥΛΗΣ  
ΙΩΑΝΝΗΣ ΠΕΛΕΚΑΝΟΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΥΠΕΥΘΥΝΟΣ

ΚΩΝΣΤΑΝΤΙΝΟΣ ΚΟΛΟΜΒΑΤΣΟΣ  
ΑΝΑΠΛΗΡΩΤΗΣ ΚΑΘΗΓΗΤΗΣ

Λαμία 2025-2026





UNIVERSITY OF  
THESSALY

SCHOOL OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE & TELECOMMUNICATIONS

# STOCK MARKET USING LLM

CHRISTOS MITSOULIS  
IOANNIS PELEKANOS

FINAL THESIS

ADVISOR

KONSTANTINOS KOLOMVATSOS  
ASSOCIATE PROFESSOR

Lamia 2025-2026



«Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις <sup>(1)</sup>, που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

1. Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί **χωρίς να τα περικλείω σε εισαγωγικά** και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάστηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.

2. Δέχομαι ότι η αυτολεξεί **παράθεση χωρίς εισαγωγικά**, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική μου εργασία.

3. Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια.

4. Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.

Ημερομηνία: 05.../..01.../2026.....

Ο – Η Δηλ.

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.»





## ΠΕΡΙΛΗΨΗ

---

Η παρούσα πτυχιακή εργασία εξετάζει τον ρόλο των σύγχρονων τεχνολογιών τεχνητής νοημοσύνης στη χρηματοοικονομική ανάλυση με χρήση μεγάλων γλωσσικών μοντέλων (Large Language Models, LLM) για την επεξεργασία δεδομένων και κειμένων. Μέσα από την ανάπτυξη εφαρμογών, παρουσιάζεται πώς οι τεχνολογίες LLM μπορούν να υποστηρίξουν την κατανόηση και αξιολόγηση χρηματοοικονομικών πληροφοριών. Τα αποτελέσματα αναδεικνύουν τόσο τις δυνατότητες όσο και τους περιορισμούς αυτών των τεχνολογιών, επιβεβαιώνοντας ότι μπορούν να λειτουργήσουν συμπληρωματικά στα παραδοσιακά εργαλεία ανάλυσης και να βελτιώσουν τη λήψη αποφάσεων σε πολύπλοκα περιβάλλοντα.

## ABSTRACT

---

The present thesis examines the role of modern artificial intelligence technologies in financial analysis using Large Language Models (LLM) for data and text processing. Through the development of applications, it demonstrates how LLM can support the understanding and evaluation of financial information. The results highlight both the capabilities and the limitations of these technologies, confirming that they can function complementarily with traditional analytical tools and enhance decision-making in complex environments.

## Table of Contents

---

ΠΕΡΙΛΗΨΗ .....	I
ABSTRACT .....	II
<b>ΚΕΦΑΛΑΙΟ 1 ΕΙΣΑΓΩΓΗ .....</b>	<b>0</b>
<b>ΚΕΦΑΛΑΙΟ 2 ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ.....</b>	<b>2</b>
<b>(2.1 ΓΛΩΣΣΙΚΗ ΜΟΝΤΕΛΟΠΟΙΗΣΗ).....</b>	<b>2</b>
<b>(2.1 LLM).....</b>	<b>3</b>
(2.1.A ΑΡΧΙΤΕΚΤΟΝΙΚΗ LLM).....	3
(2.1.B ΒΑΣΙΚΑ ΣΥΣΤΑΤΙΚΑ LLM) .....	5
(2.1.Γ ΙΚΑΝΟΤΗΤΕΣ LLM).....	8
(2.1.Δ ΠΡΟΚΛΗΣΕΙΣ LLM) .....	9
(2.1.Ε ΟΙΚΟΓΕΝΕΙΕΣ LLM).....	10
<b>ΚΕΦΑΛΑΙΟ 3 ΕΦΑΡΜΟΓΕΣ LLM ΓΙΑ STOCK MARKET.....</b>	<b>11</b>
<b>(3.1 ΜΕΘΟΔΟΛΟΓΙΑ ΥΛΟΠΟΙΗΣΗΣ ΕΦΑΡΜΟΓΩΝ STOCK MARKET) .....</b>	<b>11</b>
(3.1.A STREAMLIT).....	13
(3.1.B YFINANCE) .....	13
(3.1.Γ FINVIZFINANCE).....	14
(3.1.Δ OLLAMA) .....	15
<b>(3.2 ΑΝΑΚΤΗΣΗ ΙΣΤΟΡΙΚΩΝ ΔΕΔΟΜΕΝΩΝ STOCK MARKET) .....</b>	<b>17</b>
<b>(3.3 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ) .....</b>	<b>21</b>
(3.3.A ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΣΤΟ ΧΡΗΜΑΤΙΣΤΗΡΙΟ).....	21
(3.3.B ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΓΙΑ STOCK MARKET) .....	22
(3.3.Γ ΟΠΤΙΚΟΠΟΙΗΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ).....	27
<b>(3.4 ΠΡΟΒΛΕΨΗ ΑΠΟΔΟΣΗΣ ΜΕΤΟΧΗΣ) .....</b>	<b>29</b>
(3.4.A ΠΡΟΕΤΟΙΜΑΣΙΑ ΔΕΔΟΜΕΝΩΝ) .....	29
(3.4.B ΕΚΠΑΙΔΕΥΣΗ ΑΛΓΟΡΙΘΜΩΝ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ).....	33
(3.4.Γ ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΒΛΕΨΗΣ) .....	37
<b>(3.5 ΣΥΝΟΨΗ &amp; ΕΞΑΓΩΓΗ ΔΕΔΟΜΕΝΩΝ ΟΙΚΟΝΟΜΙΚΩΝ ΑΝΑΦΟΡΩΝ) .....</b>	<b>47</b>
(3.5.A ΤΕΧΝΙΚΗ ΠΡΟΣΕΓΓΙΣΗ) .....	47
(3.5.B ΠΕΡΙΓΡΑΦΗ ΕΦΑΡΜΟΓΗΣ).....	48
(3.5.Γ ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ).....	53
<b>(3.6 CHATBOT &amp; PROMPT ENGINEERING) .....</b>	<b>59</b>
(3.6.A ΕΙΣΑΓΩΓΗ) .....	59
(3.6.B ΠΕΡΙΓΡΑΦΗ ΕΦΑΡΜΟΓΗΣ CHATBOT).....	61
(3.6.Γ ΠΑΡΑΔΕΙΓΜΑΤΑ ΕΦΑΡΜΟΓΗΣ CHATBOT & PROMPT ENGINEERING).....	63
<b>(3.7 RED FLAGGING ΓΙΑ ΔΙΑΧΕΙΡΙΣΗ ΚΙΝΔΥΝΩΝ).....</b>	<b>76</b>
(3.7.A ΠΕΡΙΓΡΑΦΗ ΕΦΑΡΜΟΓΗΣ RED FLAGGING) .....	77
(3.7.B ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ RED FLAGGING) .....	80
<b>ΚΕΦΑΛΑΙΟ 4 ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ .....</b>	<b>83</b>

( 4.1 ΣΥΖΗΤΗΣΗ ΕΡΓΑΛΕΙΩΝ ΕΦΑΡΜΟΓΩΝ) .....	83
(4.2 ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΠΡΟΒΛΕΨΗΣ ΑΠΟΔΟΣΗΣ ΜΕΤΟΧΗΣ) .....	84
( 4.3 ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΣΥΝΟΨΗΣ & ΑΝΑΛΥΣΗΣ ΑΝΑΦΟΡΩΝ) .....	86
(4.4 ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ CHATBOT) .....	88
( 4.5 ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ RED FLAGGING) .....	91
<b><u>ΚΕΦΑΛΑΙΟ 5 ΣΥΜΠΕΡΑΣΜΑΤΑ.....</u></b>	<b>93</b>
<b><u>ΒΙΒΛΙΟΓΡΑΦΙΑ.....</u></b>	<b>94</b>



# ΚΕΦΑΛΑΙΟ 1 Εισαγωγή

---

Η συνεχής πρόοδος της Τεχνητής Νοημοσύνης (Artificial Intelligence, AI) αποτελεί έναν από τους σημαντικότερους θεμελιώδεις παράγοντες τεχνολογικής εξέλιξης στη σύγχρονη ψηφιακή πραγματικότητα. Ιδιαίτερα τα τελευταία χρόνια, η ανάπτυξη μοντέλων βαθιάς μάθησης έχει επιτρέψει την αυτοματοποίηση και την υποστήριξη σύνθετων γνωστικών διεργασιών, όπως η ανάλυση μεγάλων όγκων δεδομένων, η εξαγωγή γνώσης και η λήψη αποφάσεων σε περιβάλλοντα υψηλής πολυπλοκότητας. Στο πλαίσιο αυτό, τα μεγάλα γλωσσικά μοντέλα (Large Language Models, LLM) έχουν αναδειχθεί ως μία από τις πιο καινοτόμες και επιδραστικές τεχνολογίες, καθώς συνδυάζουν υπολογιστική ισχύ με προηγμένες δυνατότητες επεξεργασίας φυσικής γλώσσας.

Τα LLM εκπαιδεύονται σε τεράστιες συλλογές κειμένων και είναι σε θέση να κατανοούν το νόημα, το πλαίσιο και τη δομή της ανθρώπινης γλώσσας, παράγοντας απαντήσεις που προσεγγίζουν την ανθρώπινη σκέψη. Οι δυνατότητές τους εκτείνονται από την απλή σύννοψη κειμένων έως τη σύνθετη ανάλυση πληροφοριών, την εξαγωγή συμπερασμάτων και την παροχή επεξηγηματικών απαντήσεων. Η ευελιξία αυτή καθιστά τα LLM ιδιαίτερα ελκυστικά για εφαρμογές σε τομείς όπου η πληροφορία είναι κατά κύριο λόγο μη δομημένη, όπως τα οικονομικά κείμενα, οι ειδήσεις και οι εταιρικές αναφορές.

Ο χρηματοοικονομικός τομέας αποτελεί ένα από τα πλέον απαιτητικά και ανταγωνιστικά περιβάλλοντα για την εφαρμογή τεχνολογιών τεχνητής νοημοσύνης. Οι χρηματιστηριακές αγορές χαρακτηρίζονται από έντονη μεταβλητότητα, υψηλό βαθμό αβεβαιότητας και συνεχή παραγωγή δεδομένων από ετερογενείς πηγές. Οι επενδυτές και οι αναλυτές καλούνται να παρακολουθούν σε πραγματικό χρόνο τις εξελίξεις της αγοράς, να αξιολογούν οικονομικές αναφορές, να ερμηνεύουν ειδήσεις και να λαμβάνουν αποφάσεις υπό χρονική πίεση. Η παραδοσιακή ανάλυση, η οποία βασίζεται αποκλειστικά σε αριθμητικά δεδομένα ή σε ανθρώπινη ερμηνεία κειμένων, συχνά αδυνατεί να ανταποκριθεί αποτελεσματικά σε αυτές τις απαιτήσεις.

Σε αυτό το περιβάλλον, τα μεγάλα γλωσσικά μοντέλα προσφέρουν νέες δυνατότητες ενίσχυσης της χρηματιστηριακής ανάλυσης. Η ικανότητά τους να επεξεργάζονται και να ερμηνεύουν ειδήσεις, οικονομικές καταστάσεις, δελτία τύπου και στρατηγικές ανακοινώσεις επιτρέπει την ταχύτερη εξαγωγή ουσιαστικών πληροφοριών και τη μείωση του γνωστικού φόρτου για τον χρήστη. Παράλληλα, τα LLM μπορούν να αξιοποιηθούν για την ανάλυση του συναισθήματος της αγοράς, την ανίχνευση υπερβολικά αισιόδοξων ή ασαφών δηλώσεων και την υποστήριξη διαδικασιών διαχείρισης κινδύνου.

Ταυτόχρονα με την ανάλυση κειμένων, η χρήση τεχνικών μηχανικής μάθησης για την επεξεργασία χρονοσειρών μετοχών έχει αποτελέσει αντικείμενο εκτεταμένης έρευνας. Αλγόριθμοι όπως τα δέντρα αποφάσεων, τα νευρωνικά δίκτυα και τα μοντέλα ενίσχυσης επιτρέπουν την αναγνώριση μοτίβων, την εκτίμηση τάσεων και την πρόβλεψη μελλοντικών μεταβολών στις τιμές των μετοχών. Ωστόσο, οι χρηματοοικονομικές χρονοσειρές χαρακτηρίζονται από υψηλό θόρυβο και μη γραμμικές σχέσεις, γεγονός που περιορίζει την ακρίβεια των προβλέψεων και καθιστά αναγκαία τη χρήση συμπληρωματικών πηγών πληροφορίας.

Ο συνδυασμός ποσοτικών δεδομένων αγοράς με ποιοτική πληροφορία που προκύπτει από την ανάλυση φυσικής γλώσσας δημιουργεί ένα υβριδικό πλαίσιο χρηματιστηριακής ανάλυσης. Μέσα από αυτή την προσέγγιση, οι αριθμητικές προβλέψεις ενισχύονται από ερμηνευτικά στοιχεία που αφορούν τις ειδήσεις, τη στρατηγική των εταιρειών και το γενικότερο οικονομικό κλίμα. Η συνέργεια μεταξύ μηχανικής μάθησης και LLM επιτρέπει μια πιο ολιστική κατανόηση της αγοράς, η οποία δεν

περιορίζεται μόνο σε ιστορικά δεδομένα αλλά λαμβάνει υπόψη και το ευρύτερο πληροφοριακό πλαίσιο.

Η παρούσα εργασία επικεντρώνεται στη μελέτη και αξιολόγηση της χρήσης μεγάλων γλωσσικών μοντέλων στον τομέα της χρηματιστηριακής ανάλυσης, με έμφαση στο μοντέλο Llama 3. Μέσα από την ανάπτυξη διαδραστικών εφαρμογών, εξετάζεται πώς τα LLM μπορούν να ενσωματωθούν σε συστήματα πρόβλεψης απόδοσης μετοχών, σύνοψης οικονομικών αναφορών, χρηματιστηριακών chatbot και μηχανισμών εντοπισμού κινδύνων (red flagging). Η ανάλυση δεν περιορίζεται μόνο στην τεχνική απόδοση των συστημάτων, αλλά επεκτείνεται και στην πρακτική τους χρησιμότητα για επενδυτές και αναλυτές.

Ιδιαίτερη σημασία δίνεται στην αξιολόγηση των πλεονεκτημάτων και των περιορισμών των LLM, όπως ο υπολογιστικός χρόνος, η αριθμητική ακρίβεια και η ανάγκη ανθρώπινης εποπτείας. Παρά την εντυπωσιακή τους απόδοση, τα μοντέλα αυτά δεν μπορούν να αντικαταστήσουν πλήρως την ανθρώπινη κρίση, αλλά λειτουργούν ως εργαλεία υποστήριξης αποφάσεων που ενισχύουν την αναλυτική διαδικασία. Η εργασία φιλοδοξεί να συμβάλει στην κατανόηση του ρόλου των LLM στον χρηματοοικονομικό τομέα και να αναδείξει τις προοπτικές και τις προκλήσεις της μελλοντικής τους αξιοποίησης.

Η εργασία οργανώνεται ως εξής: στο Κεφάλαιο 2 παρουσιάζεται το θεωρητικό υπόβαθρο που αφορά τα μεγάλα γλωσσικά μοντέλα και τις βασικές έννοιες της μηχανικής μάθησης που σχετίζονται με τη χρηματιστηριακή ανάλυση. Στο Κεφάλαιο 3 περιγράφεται η μεθοδολογία και η υλοποίηση των εφαρμογών που αναπτύχθηκαν, συμπεριλαμβανομένων των εργαλείων που χρησιμοποιήθηκαν, των πηγών δεδομένων και των αλγορίθμων πρόβλεψης. Στο Κεφάλαιο 4 γίνεται αναλυτική παρουσίαση και συζήτηση των αποτελεσμάτων για κάθε εφαρμογή, με έμφαση στην αξιολόγηση της απόδοσης και στους περιορισμούς των προσεγγίσεων. Στο Κεφάλαιο 5 συνοψίζονται τα βασικά συμπεράσματα της εργασίας και παρουσιάζονται προτάσεις για μελλοντική έρευνα και περαιτέρω βελτίωση των συστημάτων.

# ΚΕΦΑΛΑΙΟ 2 Βιβλιογραφική Επισκόπηση

## (2.1 Γλωσσική Μοντελοποίηση)

Η γλωσσική μοντελοποίηση αποτελεί μία από τις βάσεις της τεχνητής νοημοσύνης ιδιαίτερα στον τομέα της επεξεργασίας φυσικής γλώσσας (Natural Language Processing, NLP). Έχει σαν στόχο την εκπαίδευση ενός μοντέλου να κατανοεί, να προβλέπει και να παράγει ανθρώπινη γλώσσα. Ένα γλωσσικό μοντέλο στοχεύει στην κατανόηση των στατιστικών δομών και των βαθύτερων συσχετίσεων της γλώσσας, ώστε να μπορεί να αναπαράγει το κείμενο με φυσικότητα και συνοχή [1], [2].

Στα πρώτα στάδια της ανάπτυξής της, η γλωσσική μοντελοποίηση βασίστηκε σε στατιστικές μεθόδους, όπως τα N-gram μοντέλα και τα μοντέλα βασισμένα σε γραμματικούς κανόνες, προκειμένου να προβλέπει ή να παράγει λέξεις με βάση τις πιθανότητες εμφάνισής τους σε κείμενα. Αν και τα μοντέλα αυτά ήταν απλά, είχαν περιορισμούς επειδή δεν μπορούσαν να αποτυπώσουν μεγάλες ακολουθίες ή σύνθετες εξαρτήσεις στη γλώσσα [1], [3].

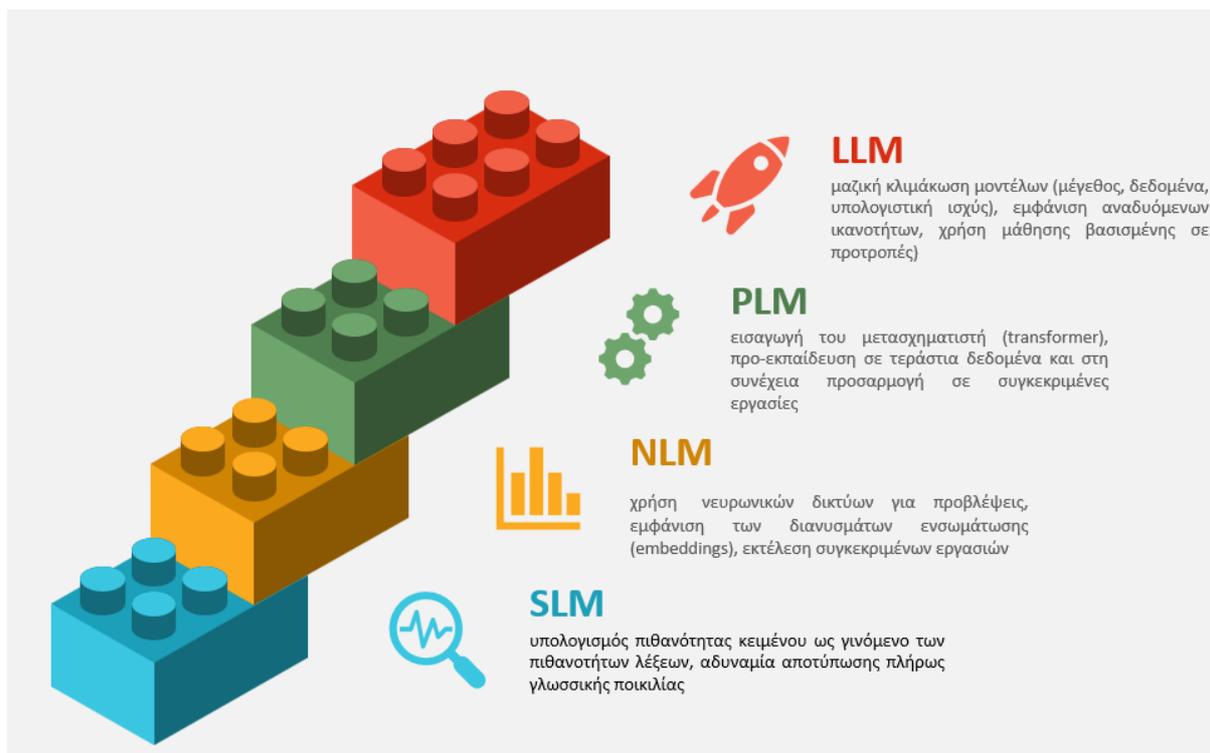
Η εξέλιξη των γλωσσικών μοντέλων ακολουθεί μια μακρά πορεία τεχνολογικής ανάπτυξης, όπου κάθε στάδιο έδωσε λύση σε προβλήματα που παρέμεναν άλυτα στο προηγούμενο, οδηγώντας τελικά στην εποχή των μεγάλων γλωσσικών μοντέλων (Large Language Models, LLM). Η αρχή αυτής της εξέλιξης (Σχήμα 1.1) εντοπίζεται στα στατιστικά γλωσσικά μοντέλα (Statistical Language Models, SLM), συνεχίζεται με τα νευρωνικά γλωσσικά μοντέλα (Neural Language Models, NLM), μετασχηματίζεται με τα γλωσσικά μοντέλα προ-εκπαίδευσης (Pre-Trained Language Models, PLM) και κορυφώνεται στα σύγχρονα μεγάλα γλωσσικά μοντέλα (LLM) [1], [2].

Τα στατιστικά γλωσσικά μοντέλα (SLM) αντιμετωπίζουν το κείμενο ως ακολουθία λέξεων και υπολογίζουν την πιθανότητά του ως το γινόμενο των πιθανοτήτων των επιμέρους λέξεων. Χρησιμοποιούνται τα μοντέλα αλυσίδας Markov, γνωστά ως n-gram μοντέλα, τα οποία εκτιμούν την πιθανότητα μιας λέξης με βάση τις προηγούμενες  $n - 1$  λέξεις. Τα μοντέλα αυτά αντιμετωπίζουν το πρόβλημα της σπανιότητας δεδομένων, δηλαδή την εμφάνιση λέξεων ή n-grams που δεν έχουν παρατηρηθεί. Παρά τη διαδεδομένη χρήση τους σε εφαρμογές επεξεργασίας φυσικής γλώσσας, τα n-gram μοντέλα παραμένουν περιορισμένα, καθώς δεν αποτυπώνουν πλήρως την ποικιλία της γλώσσας [1], [3].

Τα πρώιμα νευρωνικά γλωσσικά μοντέλα (NLM) αντιμετώπιζαν τη σπανιότητα δεδομένων χρησιμοποιώντας διανύσματα ενσωμάτωσης (embeddings) για τις λέξεις και προβλέποντας την επόμενη λέξη μέσω νευρωνικών δικτύων. Τα embeddings δημιουργούν έναν κρυφό χώρο όπου η σημασιολογική συγγένεια μπορεί να υπολογιστεί εύκολα, δίνοντας τη δυνατότητα σύγκρισης εισόδων με διαφορετική μορφή ή γλώσσα. Τα πρώιμα NLM ήταν προσαρμοσμένα σε συγκεκριμένες εργασίες [1], [2].

Τα προ-εκπαιδευμένα γλωσσικά μοντέλα (PLM), σε αντίθεση με τα πρώιμα NLM, δεν είναι εξειδικευμένα σε συγκεκριμένες εργασίες. Αυτή η γενικότητα επεκτείνεται και στον κρυφό χώρο των embeddings που εκπαιδεύονται. Η εκπαίδευση και η χρήση τους ακολουθεί το παράδειγμα προ-εκπαίδευσης (pre-training) και λεπτομερούς προσαρμογής (fine-tuning), όπου γλωσσικά μοντέλα με αναδρομικά νευρωνικά δίκτυα ή μετασχηματιστές (transformers) προ-εκπαιδεύονται σε τεράστια κείμενα για γενικές εργασίες, όπως η πρόβλεψη λέξεων, και στη συνέχεια σε συγκεκριμένες εργασίες χρησιμοποιώντας μικρές ποσότητες επισημασμένων δεδομένων (labeled data) [1], [2], [3].

Τα μεγάλα γλωσσικά μοντέλα (LLM) είναι νευρωνικά μοντέλα που βασίζονται σε transformer με δεκάδες έως εκατοντάδες δισεκατομμύρια παραμέτρους, προ-εκπαιδευμένα σε τεράστια κείμενα, όπως τα PaLM, LLaMA και GPT-4. Σε σχέση με τα προ-εκπαιδευμένα γλωσσικά μοντέλα, τα μεγάλα γλωσσικά μοντέλα είναι μεγαλύτερα, με ισχυρότερη κατανόηση και παραγωγή γλώσσας, καθώς και αναδυόμενες ικανότητες που εμφανίζονται μόνο σε μεγάλης κλίμακας μοντέλα. Αυτές περιλαμβάνουν μάθηση μέσω συμφραζομένων (in-context learning), εκτέλεση οδηγιών χωρίς παραδείγματα (instruction following) και πολύβηματική συλλογιστική (multi-step reasoning). Επιπλέον, μπορούν να ενισχυθούν με εξωτερική γνώση και εργαλεία, αλληλοεπιδρώντας με χρήστες και περιβάλλον και βελτιώνοντας συνεχώς τις επιδόσεις τους μέσω ανατροφοδότησης [1], [2].

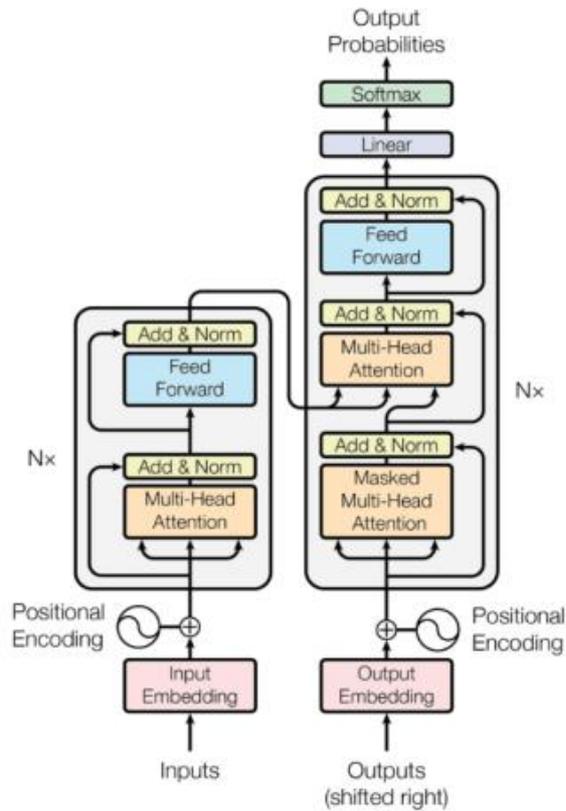


Σχήμα 2.1: Ιστορική εξέλιξη της Γλωσσικής Μοντελοποίησης

## (2.1 LLM)

### (2.1.a Αρχιτεκτονική LLM)

Η αρχιτεκτονική των μεγάλων γλωσσικών μοντέλων στηρίζεται σχεδόν αποκλειστικά στον μετασηματιστή, ένα μοντέλο που εισήγαγε μια νέα φιλοσοφία στην επεξεργασία ακολουθιών. Ο μετασηματιστής (Σχήμα 2.1) είναι μια αρχιτεκτονική νευρωνικών δικτύων για γλωσσική μοντελοποίηση και επεξεργασία ακολουθιών, η οποία βασίζεται στον μηχανισμό αυτοπροσοχής (self-attention) [1], [2]. Ουσιαστικά, αντί να επεξεργάζεται τα δεδομένα διαδοχικά όπως τα αναδρομικά νευρωνικά δίκτυα (Recurrent Neural Networks, RNN), υπολογίζει παράλληλα πόσο κάθε λέξη σε μια πρόταση επηρεάζει κάθε άλλη λέξη, επιτρέποντας την αποδοτική εκπαίδευση μεγάλων μοντέλων σε τεράστια σύνολα δεδομένων και την καλύτερη κατανόηση εκτενών εξαρτήσεων μέσα στο κείμενο [1], [2].



Σχήμα 2.2: Δομή μετασχηματιστή [1]

Στην αρχιτεκτονική μετασχηματιστή, η λειτουργία προσοχής (attention) βασίζεται σε τρεις βασικούς όρους: ερώτημα (query), κλειδί (key) και τιμή (value). Κάθε λεκτική μονάδα (token) στο κείμενο μετατρέπεται σε αυτά τα τρία διανύσματα. Η αυτοπροσοχή (self-attention) υπολογίζει πόσο πρέπει κάθε λεκτική μονάδα να «προσέξει» τις υπόλοιπες λεκτικές μονάδες της πρότασης για να δημιουργήσει μια αναπαράσταση πλούσια σε συμφραζόμενα. Συγκεκριμένα, το ερώτημα για κάθε λεκτική μονάδα συγκρίνεται με όλα τα κλειδιά για να υπολογιστούν βάρη συμβατότητας, τα οποία εφαρμόζονται στις τιμές και δίνουν μια σταθμισμένη αναπαράσταση για κάθε λεκτική μονάδα [1], [2]. Αυτό επιτρέπει στο μοντέλο να κατανοεί τις σχέσεις μεταξύ λέξεων ανεξάρτητα από τη θέση τους στην πρόταση.

Η αρχιτεκτονική μετασχηματιστή, αρχικά σχεδιασμένη για μετάφραση κειμένου, αποτελείται από έναν κωδικοποιητή (encoder) και έναν αποκωδικοποιητή (decoder), καθένας με συνήθως έξι πανομοιότυπα επίπεδα. Κάθε επίπεδο του κωδικοποιητή περιλαμβάνει δύο υποεπίπεδα: ένα πολλαπλής κεφαλής αυτοπροσοχής και ένα πλήρως συνδεδεμένο προωθητικό δίκτυο (fully connected feed-forward network). Ο αποκωδικοποιητής περιλαμβάνει αυτά τα δύο υποεπίπεδα και ένα επιπλέον υποεπίπεδο που εφαρμόζει προσοχή πάνω στην έξοδο του κωδικοποιητή [1], [2].

Τα μοντέλα μόνο με κωδικοποιητή (encoder-only models), όπως το BERT (Bidirectional Encoder Representations from Transformers), έχουν πρόσβαση σε όλες τις λέξεις μιας πρότασης και προεκπαιδεύονται συνήθως με τη μέθοδο της μάσκας λέξεων (word making method). Η μέθοδος της μάσκας λέξεων είναι μια τεχνική εκπαίδευσης γλωσσικών μοντέλων όπου τμήματα του κειμένου «κρύβονται» προσωρινά και το μοντέλο πρέπει να τα προβλέψει από το πλαίσιο. Έτσι τα μοντέλα αυτά είναι ιδανικά για κατανόηση ολόκληρου του κειμένου, όπως ταξινόμηση προτάσεων ή αναγνώριση οντοτήτων [1], [2].

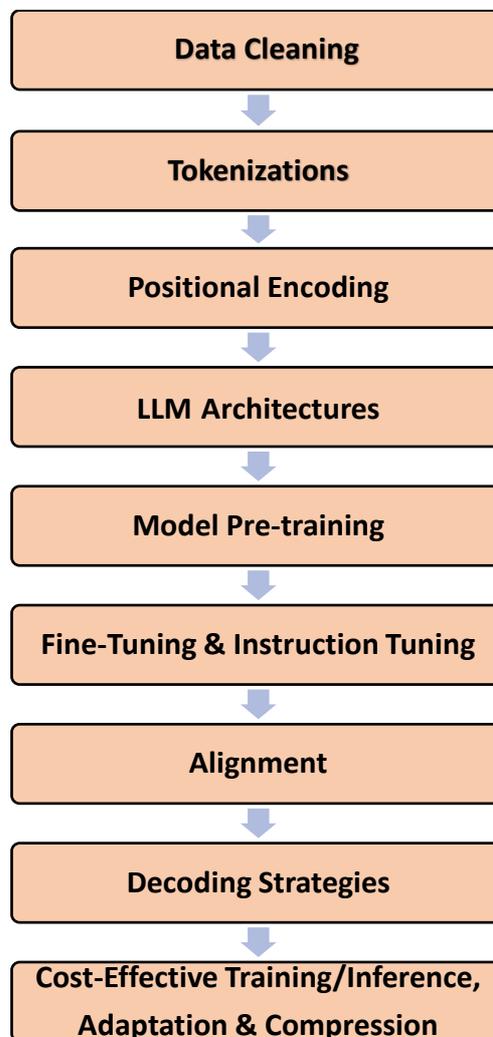
Τα μοντέλα μόνο με αποκωδικοποιητή (decoder-only models), όπως τα GPT, είναι αυτοπαλινδρομικά (auto-regressive), χρησιμοποιούν τις προηγούμενες λέξεις και προ-εκπαιδεύονται για πρόβλεψη της επόμενης λέξης, μια ικανότητα που τα καθιστά κατάλληλα για παραγωγή κειμένου [1], [2], [4].

Τέλος, τα μοντέλα κωδικοποιητή-αποκωδικοποιητή (encoder-decoder models) συνδυάζουν και τα δύο: ο κωδικοποιητής βλέπει όλο το κείμενο, ενώ ο αποκωδικοποιητής μόνο τις προηγούμενες λέξεις, και συνήθως προ-εκπαιδεύονται για πιο σύνθετες εργασίες όπως περίληψη, μετάφραση ή γενετική απάντηση (απάντηση που παράγεται αυτόματα από το μοντέλο) σε ερωτήσεις [1], [2], [3].

### (2.1.β Βασικά Συστατικά LLM)

---

Η κατασκευή ενός μεγάλου γλωσσικού μοντέλου (LLM) αποτελεί μια σύνθετη διαδικασία με πολλά στάδια που συνδυάζει δεδομένα, αρχιτεκτονική, εκπαίδευση και τεχνικές για βελτιστοποίηση όπως παρουσιάζεται στο Σχήμα 2.3 [1].



Σχήμα 2.3: Βασικά συστατικά LLM [1]

Αρχικά, συλλέγονται μεγάλα σύνολα δεδομένων που υποβάλλονται σε καθαρισμό (Data Cleaning). Ο καθαρισμός δεδομένων είναι το στάδιο της προ-επεξεργασίας όπου εντοπίζονται και διορθώνονται λάθη, ελλείψεις, ασυνέπειες, διπλότυπα και παράλογες τιμές μέσα σε ένα σύνολο δεδομένων, ώστε αυτά να γίνουν αξιόπιστα και κατάλληλα για ανάλυση ή μοντελοποίηση. Περιλαμβάνει διαδικασίες όπως αφαίρεση διπλοτύπων, διαχείριση ελλειπόντων τιμών, διόρθωση δομικών λαθών, ανίχνευση και αντιμετώπιση ακραίων τιμών (outliers), καθώς και κανονικοποίηση (normalization) μορφών (π.χ. μονάδες, ημερομηνίες, formats). Η διαδικασία είναι κρίσιμη, επειδή η ποιότητα των δεδομένων επηρεάζει άμεσα την ακρίβεια των μοντέλων και τη λήψη σωστών αποφάσεων, μειώνοντας σφάλματα, κόστος και χρόνο επεξεργασίας [1], [2].

Μετά τον καθαρισμό, ακολουθεί κατάτμηση σε λεκτικές μονάδες ή τοκενοποίηση (tokenization). Η κατάτμηση σε λεκτικές μονάδες είναι η διαδικασία κατάτμησης ενός κειμένου σε μικρότερες λεκτικές μονάδες (tokens) που μπορεί να είναι λέξεις, υπολέξεις, χαρακτήρες ή σύμβολα, ανάλογα με το μοντέλο και τον σκοπό της επεξεργασίας. Αποτελεί βασικό βήμα στην προ-επεξεργασία φυσικής γλώσσας, επειδή μετατρέπει το ακατέργαστο κείμενο σε δομημένη μορφή που μπορούν να κατανοήσουν και να επεξεργαστούν τα γλωσσικά μοντέλα. Η ποιότητα της τοκενικοποίησης επηρεάζει την ακρίβεια των εργασιών, όπως ταξινόμηση, μετάφραση, ανάλυση συναισθήματος και παραγωγή κειμένου, ενώ σε σύγχρονα μοντέλα χρησιμοποιούνται συνήθως μέθοδοι υπολεξικών λεκτικών μονάδων (subword tokenization methods) για καλύτερο χειρισμό σπάνιων λέξεων και γλωσσικής ποικιλίας [1], [3].

Στη συνέχεια εφαρμόζεται κωδικοποίηση θέσης (positional encoding). Πρόκειται για μια τεχνική που χρησιμοποιείται στα μοντέλα τύπου μετασχηματιστή για να ενσωματώσει πληροφορία σχετικά με τη σειρά των λεκτικών μονάδων μέσα σε μια πρόταση, αφού τα μοντέλα αυτά δεν έχουν εγγενή αίσθηση ακολουθίας (intrinsic sense of sequence) δηλαδή φυσική ικανότητα να κατανοεί τη σειρά των στοιχείων. Η κωδικοποίηση θέσης προσθέτει σε κάθε λεκτική μονάδα ένα διάνυσμα που αντιπροσωπεύει τη σχετική ή απόλυτη θέση του στην ακολουθία, επιτρέποντας στο μοντέλο να κατανοήσει τη σειρά, τις αποστάσεις και τις σχέσεις μεταξύ των λέξεων. Αυτό είναι κρίσιμο για εργασίες κατανόησης και παραγωγής φυσικής γλώσσας, όπου η σημασία συχνά εξαρτάται από τη θέση των λέξεων μέσα στην πρόταση [1], [2].

Τα μοντέλα βασίζονται κυρίως στην αρχιτεκτονική μετασχηματιστή, η οποία μπορεί να είναι encoder-only, decoder-only ή encoder-decoder, ανάλογα με τη χρήση τους για κατανόηση ή παραγωγή κειμένου [1].

Η προ-εκπαίδευση (model pre-training) αποτελεί το επόμενο κρίσιμο στάδιο όπου τα μοντέλα εκπαιδεύονται σε τεράστια σύνολα κειμένου με στόχο να μάθουν γενικές γλωσσικές δομές και πρότυπα, όπως συντακτικό, σημασιολογία και κατανομή λέξεων. Αυτή η προ-εκπαίδευση γίνεται συνήθως χωρίς επισημασμένα δεδομένα (self-supervised), μέσω εργασιών όπως η πρόβλεψη της επόμενης λεκτικής μονάδας ή η πρόβλεψη μάσκας σε λέξεις, επιτρέποντας στα μοντέλα να αναπτύξουν γενικευμένη ικανότητα κατανόησης και παραγωγής φυσικής γλώσσας. Η προ-εκπαίδευση δημιουργεί τη βάση πάνω στην οποία στηρίζεται η περαιτέρω προσαρμογή (fine-tuning) για συγκεκριμένες εργασίες, ενώ επιτρέπει την εμφάνιση αναδυόμενων ικανοτήτων, όπως η σύνθεση κειμένου και η κατανόηση συμφραζομένων, οι οποίες αναδεικνύουν τα LLM ως ιδιαίτερα ευέλικτα και ισχυρά σε ποικίλες γλωσσικές εφαρμογές [1], [2], [4].

Μετά την προ-εκπαίδευση, το μοντέλο υποβάλλεται σε λεπτομερή προσαρμογή (fine-tuning) και προσαρμογή σε οδηγίες (instruction tuning) για να προσαρμοστεί σε συγκεκριμένες εργασίες ή εφαρμογές, χρησιμοποιώντας δεδομένα με ετικέτες ή στόχους που καθορίζουν τη συμπεριφορά του. Μέσω της λεπτομερούς προσαρμογής, το μοντέλο μπορεί να βελτιώσει την απόδοσή του σε ειδικές εργασίες, όπως ταξινόμηση κειμένου, ερωτοαπαντήσεις ή μετάφραση. Η προσαρμογή σε οδηγίες

αποτελεί μια εξειδικευμένη μορφή προσαρμογής, όπου το μοντέλο εκπαιδεύεται ώστε να ακολουθεί οδηγίες σε φυσική γλώσσα, μαθαίνοντας να ανταποκρίνεται σε αιτήματα χρηστών με πιο σαφή και συνεπή τρόπο. Συνδυαστικά, αυτές οι διαδικασίες επιτρέπουν στα LLM να διατηρούν τις γενικές γλωσσικές τους ικανότητες που αποκτήθηκαν κατά την προ-εκπαίδευση, ενώ ταυτόχρονα γίνονται πιο χρήσιμα και μπορούν να καθοδηγηθούν και να προσαρμοστούν σε συγκεκριμένες εφαρμογές και περιβάλλοντα χρήσης [1], [3].

Στη συνέχεια, εφαρμόζεται ευθυγράμμιση (alignment), μια διαδικασία προσαρμογής της συμπεριφοράς LLM ώστε οι απαντήσεις να είναι σύμφωνες με τις ανθρώπινες αξίες, οδηγίες και προσδοκίες χρήσης. Η ευθυγράμμιση επιτυγχάνεται συνήθως μέσω συνδυασμού (instruction tuning και reinforcement learning from human feedback, RLHF), όπου οι άνθρωποι αξιολογητές καθοδηγούν το μοντέλο να προτιμά απαντήσεις που είναι ακριβείς, ασφαλείς, ηθικές και χρήσιμες. Η διαδικασία αυτή είναι κρίσιμη για τη μείωση ανεπιθύμητων συμπεριφορών, όπως η παραγωγή παραπλανητικού, τοξικού ή μεροληπτικού περιεχομένου, και συμβάλλει ώστε τα LLM να γίνουν πιο αξιόπιστα και εφαρμόσιμα σε πραγματικές συνθήκες χρήσης [1], [3].

Οι στρατηγικές αποκωδικοποίησης (decoding strategies) είναι μέθοδοι που χρησιμοποιούνται κατά την παραγωγή κειμένου από LLM για να επιλεγούν οι επόμενες λεκτικές μονάδες με τρόπο που μεγιστοποιεί την ποιότητα, τη συνοχή και τη σχετικότητα του αποτελέσματος. Συχνά περιλαμβάνουν τεχνικές όπως greedy search, beam search, sampling, top-k sampling και nucleus (top-p) sampling, καθεμία με διαφορετικό τρόπο ισορροπίας μεταξύ προβλεψιμότητας και ποικιλίας. Οι στρατηγικές αποκωδικοποίησης είναι κρίσιμες για την κατανόηση των δυνατοτήτων του μοντέλου, διότι επηρεάζουν την ακρίβεια, τη φυσικότητα και τη δημιουργικότητα των παραγόμενων απαντήσεων, ενώ επιτρέπουν τον έλεγχο πιθανών ανεπιθύμητων ή μη συνεπών εκβάσεων κατά την παραγωγή κειμένου [1], [2], [4].

Οι τεχνικές για αποδοτική εκπαίδευση, προβλεψιμότητα, προσαρμογή, συμπίεση (Cost-Effective Training, Inference, Adaptation & Compression) στοχεύουν στη μείωση του κόστους εκπαίδευσης και χρήσης των μεγάλων γλωσσικών μοντέλων (LLM). Λόγω του μεγάλου αριθμού παραμέτρων και των απαιτήσεων σε υπολογιστικούς πόρους, εφαρμόζονται μέθοδοι όπως parameter-efficient fine-tuning, knowledge distillation, quantization και pruning, που επιτρέπουν τη μείωση του μεγέθους και των απαιτήσεων μνήμης χωρίς σημαντική απώλεια απόδοσης. Παράλληλα, βελτιστοποιούνται οι διαδικασίες προβλεψιμότητας και προσαρμογής, ώστε τα μοντέλα να μπορούν να εκτελούν ειδικές εργασίες με περιορισμένους πόρους, διατηρώντας παράλληλα τις γενικές γλωσσικές τους ικανότητες. Αυτές οι στρατηγικές είναι κρίσιμες για την πρακτική εφαρμογή των LLM σε περιβάλλοντα όπου η αποδοτικότητα και το κόστος αποτελούν σημαντικούς περιορισμούς [1].

Η ανάπτυξη μεγάλων γλωσσικών μοντέλων περιλαμβάνει πολλαπλά στάδια: καθαρισμό δεδομένων για διόρθωση λαθών και ελλείψεων, τοκενικοποίηση και κωδικοποίηση θέσης για κατάλληλη επεξεργασία του κειμένου, και σχεδιασμό της αρχιτεκτονικής του μοντέλου. Ακολουθεί η προ-εκπαίδευση, όπου το μοντέλο μαθαίνει γενικές γλωσσικές δομές, ενώ με λεπτομερή προσαρμογή και προσαρμογή σε οδηγίες προσαρμόζεται σε συγκεκριμένες εργασίες και οδηγίες. Η ευθυγράμμιση εξασφαλίζει συμβατότητα με ανθρώπινες αξίες, οι στρατηγικές αποκωδικοποίησης καθορίζουν την παραγωγή κειμένου, και οι τεχνικές οικονομικής εκπαίδευσης, προσαρμογής και συμπίεσης μειώνουν τις απαιτήσεις σε πόρους, διατηρώντας την απόδοση και την ευελιξία του μοντέλου. Η ολοκληρωμένη αυτή διαδικασία εξασφαλίζει ότι τα μεγάλα γλωσσικά μοντέλα λειτουργούν αξιόπιστα, αποδοτικά και προσαρμόσιμα σε ποικίλες γλωσσικές εργασίες [1], [3].

## (2.1.γ Ικανότητες LLM)

---

Τα μεγάλα γλωσσικά μοντέλα αποτελούν μία από τις πλέον προηγμένες τεχνολογίες τεχνητής νοημοσύνης στον τομέα της φυσικής γλώσσας, και η ανάπτυξή τους έχει επιτρέψει σημαντική πρόοδο στην κατανόηση, παραγωγή και επεξεργασία γλωσσικού περιεχομένου [3],[1]. Τα μεγάλα γλωσσικά μοντέλα εκπαιδεύονται σε τεράστια σύνολα δεδομένων κειμένου, χρησιμοποιώντας τεχνικές βαθιάς μάθησης, όπως η αρχιτεκτονική των μετασχηματιστών, γεγονός που τους επιτρέπει να αντλούν πολύπλοκες σχέσεις και μοτίβα σε επίπεδο λέξεων, φράσεων και παραγράφων [2]. Μέσω αυτής της εκπαίδευσης, τα μοντέλα μπορούν να κατανοούν σύνθετα κείμενα, να εξάγουν νοήματα, να αναγνωρίζουν συμφραζόμενα και να ερμηνεύουν λεκτικές και σημασιολογικές σχέσεις, υποστηρίζοντας εφαρμογές όπως ταξινόμηση κειμένου, ανάλυση συναισθήματος και ερωτοαπαντήσεις [1].

Η παραγωγή κειμένου αποτελεί μία από τις κύριες δυνατότητες των μεγάλων γλωσσικών μοντέλων. Τα μοντέλα μπορούν να δημιουργούν συνεκτικά και δομημένα κείμενα διαφορετικών ειδών, όπως επιστημονικά άρθρα, ιστορίες, ποιήματα ή τεχνικά εγχειρίδια, προσαρμόζοντας το ύφος και τον τόνο γραφής ανάλογα με το κοινό ή τις απαιτήσεις του χρήστη [4]. Επιπλέον, υποστηρίζουν πολύγλωσση μετάφραση και παραφράσεις, καθώς και απλοποιήσεις κειμένων ώστε να γίνονται πιο προσιτά σε διαφορετικά επίπεδα αναγνωστικής ικανότητας [2]. Η δυνατότητα σύνοψης επιτρέπει τη συμπύκνωση μεγάλων εγγράφων διατηρώντας τα κύρια νοήματα, ενώ η ανάλυση περιεχομένου συμβάλλει στην εξαγωγή πληροφοριών από ποικίλα δεδομένα.

Τα μεγάλα γλωσσικά μοντέλα διαθέτουν επίσης ανεπτυγμένες ικανότητες λογικής και επεξεργασίας πληροφοριών σε πολλαπλά στάδια. Μπορούν να εκτελούν αριθμητικούς και αναλυτικούς υπολογισμούς, να εφαρμόζουν αλυσίδες συλλογισμού για την επίλυση γρίφων ή σύνθετων προβλημάτων, και να παρέχουν υποδείξεις για τη λήψη αποφάσεων [3]. Στον προγραμματισμό, μπορούν να δημιουργούν κώδικα σε διαφορετικές γλώσσες, να εντοπίζουν σφάλματα και να προτείνουν διορθώσεις, και διαθέτουν την ικανότητα να παρέχουν συμβουλές για βελτιωμένες πρακτικές ανάπτυξης λογισμικού [1].

Η ικανότητα ανάκτησης γνώσης αποτελεί επίσης σημαντικό χαρακτηριστικό των μεγάλων γλωσσικών μοντέλων. Τα μοντέλα ενσωματώνουν εκτεταμένες βάσεις γνώσεων σε πολλούς τομείς, όπως η επιστήμη, η τεχνολογία, η ιστορία και οι κοινωνικές επιστήμες, ενώ η δυνατότητα ενσωμάτωσης εργαλείων περιήγησης στο διαδίκτυο επιτρέπει την ενημέρωσή τους με τρέχοντα γεγονότα και επικαιροποιημένες πληροφορίες [4]. Επιπλέον, τα μεγάλα γλωσσικά μοντέλα μπορούν να αναλύουν σύνολα δεδομένων, να αναγνωρίζουν μοτίβα και να εξάγουν συμπεράσματα με τρόπο κατανοητό για τον άνθρωπο [2].

Η δημιουργικότητα των μεγάλων γλωσσικών μοντέλων δεν περιορίζεται μόνο στη γραπτή γλώσσα. Τα μοντέλα μπορούν να παράγουν καλλιτεχνικό και διαφημιστικό περιεχόμενο, όπως σενάρια, ποιήματα ή κείμενα μάρκετινγκ, ενώ η ικανότητα τους να χειρίζονται πολλαπλά μέσα επιτρέπει να επεξεργάζονται εικόνες και να δημιουργούν περιγραφές οπτικού περιεχομένου, υποστηρίζοντας συνεργασίες με εργαλεία τεχνητής νοημοσύνης που χειρίζονται διαφορετικούς τύπους πληροφορίας [1],[4].

Επιπλέον, τα μεγάλα γλωσσικά μοντέλα μπορούν να λειτουργούν ως συνομιλητές, δάσκαλοι ή μέντορες, να προσομοιώνουν διαφορετικούς ρόλους για εκπαιδευτικούς ή δημιουργικούς σκοπούς και να εκτελούν πολύπλοκες εντολές που δίνονται από τον χρήστη [3].

## (2.1.8 Προκλήσεις LLM)

---

Παρά τις εντυπωσιακές δυνατότητες των μεγάλων γλωσσικών μοντέλων, υπάρχουν σημαντικές προκλήσεις που σχετίζονται τόσο με την τεχνική τους λειτουργία όσο και με την κοινωνική και ηθική τους διάσταση [1], [3].

Η πρώτη σημαντική πρόκληση που αντιμετωπίζουν τα μεγάλα γλωσσικά μοντέλα αφορά τη γενίκευση και τη σταθερότητα σε διαφορετικές εργασίες. Παρά το γεγονός ότι τα μοντέλα μπορούν να επιτυγχάνουν εντυπωσιακά αποτελέσματα σε εργασίες όπως η μετάφραση, η παραγωγή κειμένου ή η ανάλυση συναισθήματος, η απόδοσή τους συχνά μειώνεται όταν καλούνται να αντιμετωπίσουν πιο εξειδικευμένα ή σύνθετα προβλήματα που απαιτούν λογική, σκέψη πολλών σταδίων ή εξαγωγή συμπερασμάτων [1], [3]. Αυτό υποδεικνύει ότι η ικανότητα των μεγάλων γλωσσικών μοντέλων να γενικεύουν πέρα από τα δεδομένα εκπαίδευσής τους παραμένει περιορισμένη, θέτοντας περιορισμούς στη χρησιμότητά τους σε πιο απαιτητικά ή εξειδικευμένα περιβάλλοντα [2].

Μια δεύτερη κρίσιμη πρόκληση αφορά την αξιοπιστία και την ακρίβεια των απαντήσεων των μοντέλων. Τα μεγάλα γλωσσικά μοντέλα έχουν την τάση να παράγουν ανακριβείς ή και εντελώς φανταστικές πληροφορίες, ένα φαινόμενο που οι ερευνητές ονομάζουν «παραισθήσεις» («hallucinations») [1], [3], [4]. Το πρόβλημα αυτό είναι ιδιαίτερα έντονο όταν τα μοντέλα καλούνται να παρέχουν πληροφορίες που δεν περιλαμβάνονται στα δεδομένα εκπαίδευσής τους ή που απαιτούν επικαιροποιημένη γνώση. Η ύπαρξη αυτών των ανακριβειών καθιστά τη χρήση τους σε κρίσιμες εφαρμογές, όπως η ιατρική, η νομική ή η οικονομική ανάλυση, πιο σύνθετη και απαιτεί πρόσθετα μέτρα επαλήθευσης των αποτελεσμάτων [2], [3].

Η τρίτη πρόκληση σχετίζεται με την εξάρτηση των μεγάλων γλωσσικών μοντέλων από τα δεδομένα εκπαίδευσής. Η ποιότητα, η ποικιλία και η κάλυψη των δεδομένων επηρεάζουν άμεσα την απόδοση και τις συμπεριφορές του μοντέλου [1], [3]. Εάν τα δεδομένα είναι μεροληπτικά, ελλιπή ή εστιάζουν μόνο σε ορισμένες γλώσσες, πολιτισμικά πλαίσια ή θεματολογίες, το μοντέλο μπορεί να εμφανίζει περιορισμένη απόδοση σε άλλες περιπτώσεις ή να ενισχύει υπάρχουσες μεροληψίες. Αυτό θέτει προκλήσεις όσον αφορά την δικαιοσύνη, την ισότητα και τη συνολική αξιοπιστία των εφαρμογών τους σε παγκόσμια κλίμακα [2].

Μια ακόμη πρόκληση είναι η έλλειψη συνεπούς λογικής και σκέψης πολλών σταδίων. Παρά την πρόοδο σε τεχνικές όπως οι αλυσίδες σκέψης (Chain-of-Thought), τα μεγάλα γλωσσικά μοντέλα δεν μπορούν να εγγυηθούν ορθά ή συνεπή αποτελέσματα όταν καλούνται να λύσουν προβλήματα που απαιτούν πολλαπλά βήματα λογικής, συγκέντρωση πληροφοριών ή αιτιολόγηση [1], [3]. Αυτή η αδυναμία περιορίζει τη χρήση τους σε πολύπλοκες εφαρμογές που απαιτούν υψηλή ακρίβεια στη σκέψη και τη λήψη αποφάσεων.

Οι ηθικές και κοινωνικές προκλήσεις είναι ιδιαίτερα σημαντικές. Τα μεγάλα γλωσσικά μοντέλα έχουν τη δυνατότητα να παράγουν περιεχόμενο που είναι μεροληπτικό, τοξικό ή παραπλανητικό, ενώ μπορεί να ενισχύουν ανισότητες λόγω πολιτισμικών και γλωσσικών περιορισμών στα δεδομένα εκπαίδευσής [1], [3], [4]. Η χρήση τους σε πραγματικές συνθήκες απαιτεί προσεκτική διαχείριση, μηχανισμούς ελέγχου και στρατηγικές ευθυγράμμισης ώστε να ελαχιστοποιούνται οι αρνητικές επιπτώσεις και να διασφαλίζεται η υπεύθυνη εφαρμογή τους.

Η αντιμετώπιση όλων αυτών των προκλήσεων απαιτεί συνδυασμένες προσεγγίσεις που περιλαμβάνουν βελτίωση των δεδομένων εκπαίδευσής, τεχνικές ευθυγράμμισης, ελέγχους ποιότητας και συνεχείς αξιολογήσεις της συμπεριφοράς των μοντέλων [1], [3].

## (2.1.ε Οικογένειες LLM)

---

Οι πιο γνωστές οικογένειες μεγάλων γλωσσικών μοντέλων μπορούν να ταξινομηθούν σε τρεις βασικές κατηγορίες: GPT, LLaMA και PaLM [1]. Κάθε μία από αυτές χαρακτηρίζεται από συγκεκριμένες αρχιτεκτονικές επιλογές, στρατηγικές εκπαίδευσης και προσανατολισμό σε διαφορετικά είδη εφαρμογών, ενώ παρουσιάζει διακριτά χαρακτηριστικά όσον αφορά το μέγεθος, την απόδοση και την αξιοποίηση των δεδομένων εκπαίδευσης. Η γνώση αυτών των διαφορών είναι κρίσιμη για την επιλογή του κατάλληλου μοντέλου ανάλογα με τις ανάγκες μιας εφαρμογής, είτε αυτή αφορά επεξεργασία φυσικής γλώσσας, δημιουργία περιεχομένου, ανάλυση δεδομένων ή υποστήριξη λήψης αποφάσεων [1], [2].

Η οικογένεια GPT (Generative Pre-trained Transformer) αποτελεί ίσως την πιο αναγνωρίσιμη κατηγορία LLM και έχει επηρεάσει σημαντικά την ανάπτυξη συστημάτων συνομιλίας και δημιουργικών εργαλείων [1,2]. Τα μοντέλα GPT βασίζονται στην αρχιτεκτονική μετασχηματιστών και εκπαιδεύονται σε τεράστια πολυγλωσσικά σώματα κειμένων μέσω διαδικασίας προ-εκπαίδευσης και λεπτομερούς προσαρμογής. Αυτή η προσέγγιση επιτρέπει στα μοντέλα να κατανοούν το συμφραζόμενο και να παράγουν συνεκτικό και συνεχή λόγο, ενώ υποστηρίζει εξειδικευμένες λειτουργίες όπως ερωταποκρίσεις, σύνοψη, παραφράσεις και δημιουργική γραφή. Η ικανότητα των GPT να προσαρμόζονται σε διαφορετικά στυλ και τόνους γραφής καθιστά αυτά τα μοντέλα ιδιαίτερα χρήσιμα για εφαρμογές επικοινωνίας, εκπαίδευσης και παραγωγής περιεχομένου [1], [3].

Η οικογένεια LLaMA (Large Language Model Meta AI), αναπτυγμένη από την Meta, εστιάζει στη βελτιστοποίηση της απόδοσης σε μικρότερο μέγεθος παραμέτρων, χωρίς σημαντική μείωση της ποιότητας γλωσσικής κατανόησης [1], [4]. Τα μοντέλα LLaMA σχεδιάστηκαν για να προσφέρουν υψηλή αποδοτικότητα και δυνατότητα κλιμάκωσης, γεγονός που τα καθιστά ιδανικά για περιβάλλοντα όπου οι υπολογιστικοί πόροι είναι περιορισμένοι. Παράλληλα, η Meta έχει ενσωματώσει τεχνικές που επιτρέπουν στα μοντέλα να διατηρούν ακρίβεια και ευαισθησία στο συμφραζόμενο ακόμη και με μειωμένο μέγεθος, καθιστώντας τα κατάλληλα για επιχειρησιακές εφαρμογές, προσωπικούς βοηθούς και ενσωμάτωση σε πολυτροπικά συστήματα (multimodal systems) [4].

Η οικογένεια PaLM (Pathways Language Model) από την Google ξεχωρίζει για την ικανότητά της να διαχειρίζεται τεράστια σύνολα δεδομένων και παραμέτρων, επιτρέποντας εκτεταμένη γενίκευση και υψηλή απόδοση σε σύνθετες γλωσσικές εργασίες [1]. Το PaLM εισάγει καινοτομίες στην αρχιτεκτονική εκπαίδευσης, όπως η χρήση πολλαπλών διαδρομών, που επιτρέπει στα μοντέλα να κατανέμουν δυναμικά την επεξεργαστική ισχύ για διαφορετικές εργασίες, βελτιώνοντας την ευελιξία και την αποδοτικότητα. Αυτή η προσέγγιση επιτρέπει στα μοντέλα PaLM να εκτελούν πολύπλοκες εργασίες, όπως πολυγλωσσική μετάφραση, εξειδικευμένη ανάλυση δεδομένων και σύνθετη λογική συλλογισμού, ενώ υποστηρίζει υψηλή ακρίβεια σε εφαρμογές επιστήμης δεδομένων και μηχανικής μάθησης [1], [2].

Κοινό χαρακτηριστικό όλων των παραπάνω οικογενειών είναι η εξάρτησή τους από μεγάλες βάσεις δεδομένων κειμένου, η χρήση αρχιτεκτονικής μετασχηματιστών και η δυνατότητα προσαρμογής σε ποικίλες εφαρμογές, από απλές ερωταποκρίσεις έως σύνθετη ανάλυση δεδομένων [1], [2]. Ωστόσο, κάθε οικογένεια διαφοροποιείται στην προσέγγιση εκπαίδευσης, στο μέγεθος των παραμέτρων, στην αποδοτικότητα και στις επιδόσεις σε συγκεκριμένες εργασίες, γεγονός που επηρεάζει την επιλογή τους ανάλογα με το στόχο χρήσης και τους διαθέσιμους πόρους. Η κατανόηση των διαφορών αυτών είναι κρίσιμη για τους ερευνητές και τους επαγγελματίες της τεχνητής νοημοσύνης, καθώς επηρεάζει την ακρίβεια, την ταχύτητα και την αποτελεσματικότητα των εφαρμογών, ενώ λαμβάνει υπόψη και ηθικές παραμέτρους, όπως η μεροληψία των δεδομένων και η υπεύθυνη χρήση των μοντέλων [1], [2], [3],[ 4].

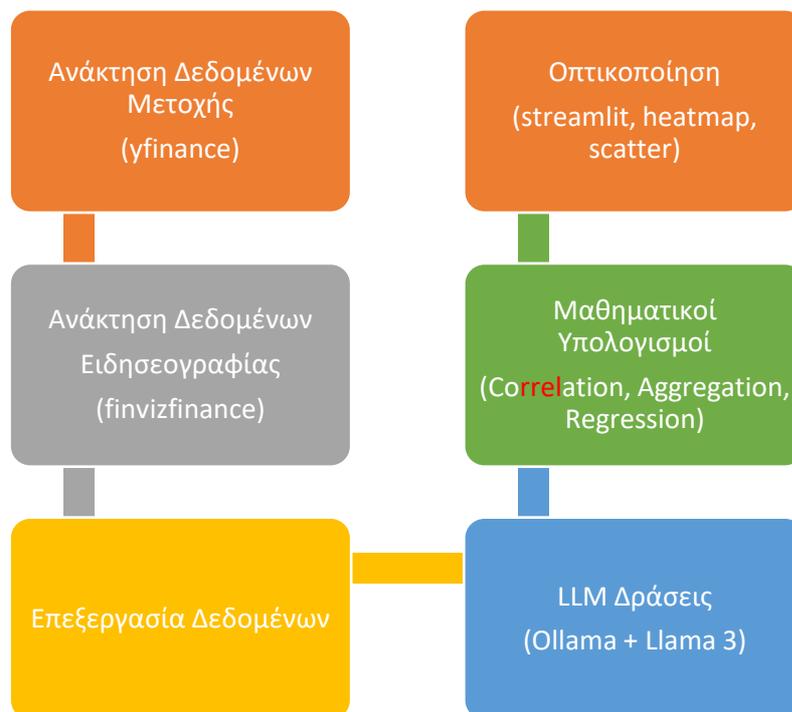
# ΚΕΦΑΛΑΙΟ 3 Εφαρμογές LLM για Stock Market

## (3.1 Μεθοδολογία Υλοποίησης Εφαρμογών Stock Market)

Στην παρούσα εργασία υλοποιήθηκαν εφαρμογές LLM για Stock Market που βασίστηκαν στη γλώσσα προγραμματισμού Python, αξιοποιώντας συγκεκριμένα εργαλεία και βιβλιοθήκες για κάθε στάδιο της διαδικασίας. Η ανάκτηση ιστορικών δεδομένων πραγματοποιήθηκε με τη χρήση των βιβλιοθηκών `yfinance` και `finvizfinance`, οι οποίες παρέχουν εύκολη πρόσβαση σε χρηματιστηριακές πληροφορίες και οικονομικά στοιχεία. Στη συνέχεια, τα δεδομένα υποβλήθηκαν σε επεξεργασία και καθαρισμό, προκειμένου να διασφαλιστεί η πληρότητα και η ακρίβεια των πληροφοριών για την περαιτέρω ανάλυση.

Οι δράσεις των μεγάλων γλωσσικών μοντέλων (LLM) εκτελέστηκαν μέσω της πλατφόρμας Ollama, χρησιμοποιώντας το μοντέλο Llama 3, το οποίο επέτρεψε την ανάλυση και πρόβλεψη τάσεων στην αγορά, καθώς και την αξιολόγηση σύνθετων δεδομένων. Παράλληλα, οι μαθηματικοί υπολογισμοί πραγματοποιήθηκαν με κατάλληλες συναρτήσεις Python.

Η οπτικοποίηση και η διάδραση με τον χρήστη υποστηρίχθηκαν από τη βιβλιοθήκη streamlit, σε συνδυασμό με επιπλέον εργαλεία όπως η seaborn για τη δημιουργία heatmaps και scatter plots. Το σύνολο της μεθοδολογίας παρουσιάζεται συνοπτικά στο Σχήμα 3.1, το οποίο απεικονίζει τη ροή εργασιών από την ανάκτηση δεδομένων έως την ανάλυση και παρουσίαση των αποτελεσμάτων.



Σχήμα 3.1: Μεθοδολογία και εργαλεία υλοποίησης εφαρμογών

Οι εφαρμογές για το χρηματιστήριο αναπτύχθηκαν με χρήση της γλώσσας προγραμματισμού Python και οι λεπτομέρειες των εργαλείων και βιβλιοθηκών που χρησιμοποιήθηκαν παρουσιάζονται στον Πίνακα 3.1.

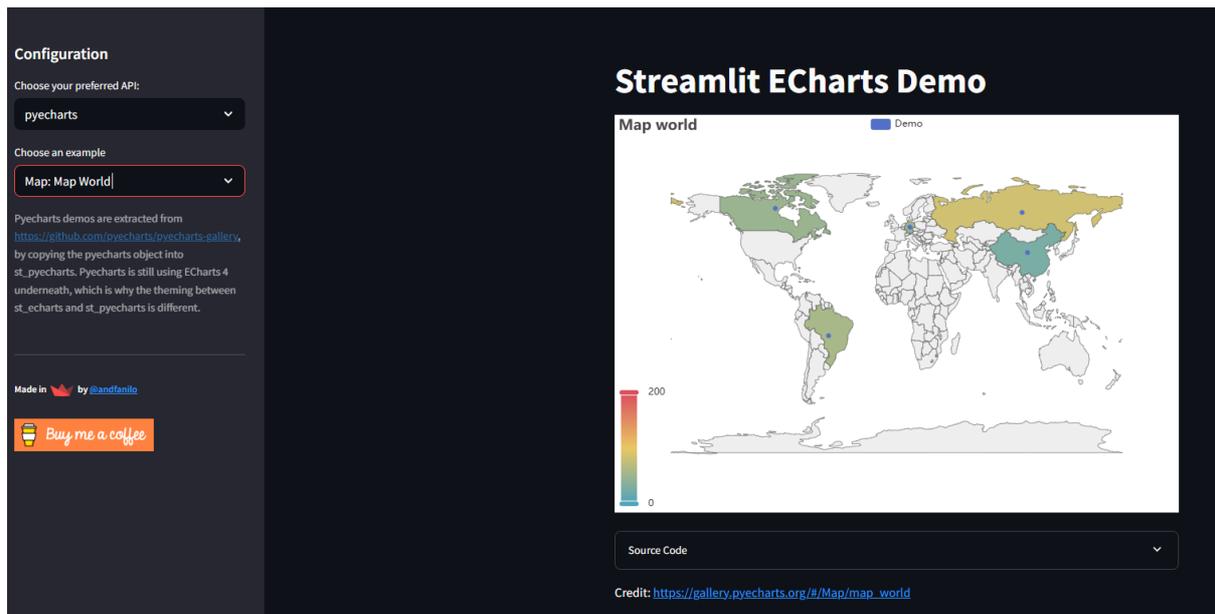
**Πίνακας 3.1:** Λίστα υλοποιημένων εφαρμογών Stock Market

No	Εφαρμογή Stock Market	Εργαλεία	Σκοπός	Παράγραφος
1	Ανάκτηση ιστορικών δεδομένων	yfinance finvizfinance streamlit	Ανάκτηση οικονομικών δεδομένων και ειδησεογραφίας	3.2
2	Κατηγοριοποίηση συναισθήματος	yfinance finvizfinance streamlit Llama 3	Κατηγοριοποίηση συναισθήματος με αιτιολόγηση σε σχέση με την ειδησιογραφία που αφορά τη μετοχή με χρήση των δεδομένων της εφαρμογής 1	3.3.β
3	Οπτικοποίηση κατηγοριοποιημένου συναισθήματος με heatmap	yfinance finvizfinance streamlit Llama 3 seaborn	Εξειδικευμένη οπτικοποίηση κατηγοριοποίησης συναισθήματος με χρήση heatmap σε συνέχεια της εφαρμογής 3	3.3.γ
5	Πρόβλεψη απόδοσης μετοχής	yfinance streamlit Llama 3 sklearn	Προβλέψεις σχετικά με την τιμή μετοχής σε συνέχεια της εφαρμογής 2	3.4.
6	Σύνοψη & Εξαγωγή Δεδομένων Οικονομικών Αναφορών	streamlit Llama 3 PyMuPDF	Σύνοψη & Εξαγωγή Δεδομένων αναφορών με οικονομικά δεδομένα και στρατηγικές	3.5
7	Chatbot & Prompt Engineering	streamlit Llama 3	Εφαρμογές chatbot σε συνδυασμό με prompt engineering	3.6
8	Red Flagging για Διαχείριση Κινδύνων	streamlit Llama 3	Εφαρμογή red flag για διαχείριση κινδύνων και υποβοήθηση λήψης αποφάσεων	3.7

### (3.1.α Streamlit)

Το streamlit (<https://streamlit.io>) αποτελεί μια ανοιχτού κώδικα βιβλιοθήκη της Python που επιτρέπει τη γρήγορη ανάπτυξη διαδραστικών web εφαρμογών για ανάλυση δεδομένων, μηχανική μάθηση και οπτικοποίηση. Η χρήση του δεν απαιτεί εξειδικευμένες γνώσεις web development, όπως HTML, CSS ή JavaScript, καθώς η δημιουργία διεπαφών πραγματοποιείται αποκλειστικά μέσω κώδικα Python.

Εκτελείται με την εντολή: `streamlit run το_αρχείο.py`



Σχήμα 3.2: Παράδειγμα εφαρμογής streamlit

### (3.1.β yfinance)

Το yfinance (<https://pypi.org/project/yfinance>) αποτελεί μια βιβλιοθήκη της Python που επιτρέπει την εύκολη και γρήγορη ανάκτηση χρηματιστηριακών δεδομένων από την πλατφόρμα Yahoo Finance. Μέσω αυτής της βιβλιοθήκης, οι χρήστες μπορούν να λαμβάνουν ιστορικά δεδομένα τιμών μετοχών, πληροφορίες για τον όγκο συναλλαγών και άλλα σημαντικά χρηματοοικονομικά στοιχεία.

Python Κώδικας για χρήση yfinance

```
import yfinance as yf
msft = yf.Ticker("MSFT") # MSFT stock ticker
print(msft.history(period="1y"))
```

Ένα ticker αναφέρεται σε έναν μοναδικό κωδικό που χρησιμοποιείται για την αναγνώριση μιας μετοχής σε χρηματιστηριακές πλατφόρμες. Παρακολουθεί τις συνεχείς αλλαγές στην τιμή του τίτλου κατά τη διάρκεια της συνεδρίασης συναλλαγών, εμφανίζοντας τόσο τις ανοδικές όσο και τις

καθοδικές κινήσεις («ticks»). Επιπλέον, παρέχει κρίσιμες πληροφορίες όπως ο όγκος συναλλαγών, οι οποίες χρησιμοποιούνται από επενδυτές και αναλυτές για την αξιολόγηση της τρέχουσας δραστηριότητας της αγοράς και του ενδιαφέροντος για τον συγκεκριμένο τίτλο.

Παράδειγμα stock ticker

Symbol ^	Company Name	Industry	Market Cap
MSFT	Microsoft Corporation	Software - Infrastructure	3,650.05B

Με την `.history()` ανακτώνται τα παρακάτω δεδομένα:

- Open: Τιμή ανοίγματος
- High: Υψηλότερη τιμή
- Low: Χαμηλότερη τιμή
- Close: Τιμή κλεισίματος
- Volume: Όγκος συναλλαγών
- Dividends: Μερίσματα (αν υπάρχουν)
- Stock Splits: Πληροφορίες για splits (αν υπάρχουν)

Πίνακας 3.2: Τυπική απάντηση δεδομένων `yfinance`

Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
26-06-2025	210.559998	212.270004	208.419998	211.389999	53200000	0	0
27-06-2025	211.130005	213.330002	210.110001	212.369995	48000000	0	0
28-06-2025	212.000000	213.949997	211.500000	213.070007	45500000	0	0
01-07-2025	213.500000	214.389999	212.029999	213.649994	49000000	0	0
02-07-2025	213.750000	215.000000	213.250000	214.910004	51000000	0	0

### (3.1.γ `finvizfinance`)

Το `finvizfinance` (<https://pypi.org/project/finvizfinance>) αποτελεί μια βιβλιοθήκη της Python που επιτρέπει την πρόσβαση σε χρηματιστηριακά δεδομένα από τον ιστότοπο `Finviz.com` μέσω εξαγωγής δεδομένων από ιστοσελίδες (`web scraping`). Παρέχει εργαλεία για οπτική παρουσίαση και ανάλυση των μετοχών, όπως `heatmaps`, ειδησεογραφία, θεμελιώδη και τεχνικά διαγράμματα, καθώς και δυνατότητα χρηματοοικονομικού ελέγχου βάσει συγκεκριμένων κριτηρίων.

Η βιβλιοθήκη είναι ιδανική για φιλτράρισμα μετοχών (`stock screening`), γρήγορη οπτική κατανόηση της αγοράς και ενημέρωση μέσω ειδήσεων, ενώ δεν προορίζεται για πολύπλοκη προγραμματιστική ανάλυση δεδομένων μέσω κώδικα.

Python Κώδικας για χρήση `finvizfinance`

```
from finvizfinance.quote import finvizfinance
stock = finvizfinance('AAPL')
print(stock.TickerOverview)
```

### (3.1.8 Ollama)

---

Το Ollama είναι μια πλατφόρμα που επιτρέπει στους χρήστες να αλληλοεπιδρούν μέσω μιας διεπαφής προγραμματισμού εφαρμογών (Application Programming Interface, API). Είναι ένα ισχυρό εργαλείο για τη δημιουργία κειμένου, την απάντηση ερωτήσεων και την εκτέλεση σύνθετων εργασιών επεξεργασίας φυσικής γλώσσας. Το Ollama παρέχει πρόσβαση σε διάφορα βελτιωμένα LLM, επιτρέποντας στους προγραμματιστές και τους ερευνητές να ενσωματώνουν εξελεγχμένες δυνατότητες κατανόησης και δημιουργίας γλώσσας στις εφαρμογές τους, όπως ρομπότ συνομιλίας (chatbots), εργαλεία δημιουργίας περιεχομένου και ερευνητικά έργα. Με το εύχρηστο API του, το Ollama βελτιστοποιεί τη διαδικασία αξιοποίησης προηγμένων μοντέλων τεχνητής νοημοσύνης, καθιστώντας την προσβάσιμη σε ένα ευρύ φάσμα χρηστών σε διαφορετικούς τομείς.



## Get up and running with large language models.

Run [DeepSeek-R1](#), [Qwen 3](#), [Llama 3.3](#), [Qwen 2.5-VL](#), [Gemma 3](#), and other models, locally.

**Σχήμα 3.3:** Ollama (<https://ollama.com>)

Η πλατφόρμα Ollama υποστηρίζει μια ευρεία γκάμα μοντέλων μεγάλων γλωσσικών συστημάτων, επιτρέποντας στους χρήστες να επιλέγουν ανάλογα με τις ανάγκες τους σε υπολογιστικούς πόρους, ακρίβεια και ειδικευση. Κορυφαίο ανάμεσά τους είναι το Llama 3, ένα θεμελιώδες μοντέλο (foundation model) με μεγάλες δυνατότητες κατανόησης και παραγωγής γλώσσας, διαθέσιμο σε εκδόσεις με παραμέτρους έως και 405 δισεκατομμύρια και με context έως 128K tokens. Η Ollama υποστηρίζει επίσης το Gemma 2, ελαφρύτερο μοντέλο κατάλληλο για εργασίες με περιορισμένους πόρους, καθώς και το Mistral-Nemo, το οποίο επιτρέπει μεγαλύτερο περιεχόμενο και σύνθετη συλλογιστική. Για προγραμματιστικές εφαρμογές, προσφέρεται το Code Llama, εξειδικευμένο στην αυτόματη δημιουργία και ανάλυση κώδικα, ενώ μοντέλα όπως το Phi 3 και οι εκδόσεις "lite" (TinyLlama, Llama 3.2) παρέχουν ελαφρύτερες επιλογές για γρήγορη δημιουργία πρωτοτύπου ή χρήση σε περιορισμένο hardware. Η διαθεσιμότητα αυτών των διαφορετικών μοντέλων επιτρέπει την προσαρμογή της πλατφόρμας σε ποικίλα σενάρια χρήσης, από πολύπλοκες εφαρμογές γενικής χρήσης και πολυμορφικές εργασίες, έως εξειδικευμένες εργασίες προγραμματισμού και εφαρμογές με περιορισμένους πόρους, εξασφαλίζοντας παράλληλα υψηλή ευελιξία και αποδοτικότητα.

Για τη συγκεκριμένη εργασία έχει επιλεγεί το Llama 3.

Η πλατφόρμα Ollama αξιοποιεί τα μοντέλα Llama 3 της Meta για την ανάπτυξη εφαρμογών τεχνητής νοημοσύνης, παρέχοντας στους χρήστες μια φιλική και εύχρηστη διεπαφή για την εκτέλεση ποικίλων γλωσσικών εργασιών. Μέσω της Ollama, τα θεμελιώδη μοντέλα Llama 3 μπορούν να ενσωματωθούν άμεσα σε εφαρμογές που απαιτούν φυσική γλώσσα, όπως συνομιλητικά συστήματα, ανάλυση κειμένου, δημιουργία περιεχομένου και υποστήριξη λήψης αποφάσεων, χωρίς να απαιτείται εξειδικευμένη γνώση στην εκπαίδευση ή τη διαχείριση μεγάλων μοντέλων. Η πλατφόρμα αξιοποιεί την ισχύ των Llama 3 για να προσφέρει συνεκτικές, ακριβείς και πολύ λειτουργικές λύσεις σε πραγματικό χρόνο, επιτρέποντας την εκμετάλλευση των δυνατοτήτων των μεγάλων γλωσσικών μοντέλων σε ποικίλα σενάρια χρήσης [4].

Το Llama 3 αποτελεί μια σύγχρονη οικογένεια θεμελιωδών γλωσσικών μοντέλων που έχουν σχεδιαστεί για ευρεία γλωσσική κατανόηση, παραγωγή κειμένου, συλλογιστική (reasoning), κωδικοποίηση (coding) και υποστήριξη εργαλείων (tool use). Στην αξιολόγηση που πραγματοποίησε η Meta, τα μοντέλα Llama 3 επιδεικνύουν απόδοση συγκρίσιμη με τα κορυφαία μοντέλα της αγοράς, όπως το GPT-4, σε πληθώρα εργασιών επεξεργασίας φυσικής γλώσσας, από την κατανόηση σύνθετων κειμένων μέχρι τη δημιουργία συνεκτικού και λογικά δομημένου περιεχομένου [4].

Ένα από τα σημαντικά πλεονεκτήματα του Llama 3 είναι η ικανότητά του να διαχειρίζεται εξαιρετικά μεγάλα περιεχόμενα, έως και 128000 tokens, καθιστώντας το κατάλληλο για πολύ μεγάλου μήκους έγγραφα ή εκτεταμένες συνομιλίες, όπου η διατήρηση συνοχής και πληροφορίας είναι κρίσιμη. Σε σύγκριση, τα μοντέλα GPT-4 διαθέτουν όρια περιεχομένου περίπου 32000 tokens στην μεγαλύτερη έκδοσή τους, γεγονός που περιορίζει την αποτελεσματικότητα σε υπερμεγέθη έγγραφα.

Η αξιολόγηση [4] έδειξε ότι το Llama 3 επιτυγχάνει υψηλή ποιότητα στην αξιολόγηση συλλογιστικής (benchmarks reasoning), εκτέλεση οδηγιών (instruction following) και εργασίες κοινής λογικής (common-sense tasks), πλησιάζοντας ή και σε ορισμένες περιπτώσεις υπερβαίνοντας την απόδοση του GPT-4. Παράλληλα, η Meta δημοσιεύει προ-εκπαιδευμένες (pre-trained) και μετα-εκπαιδευμένες (post-trained) εκδόσεις του Llama 3, καθώς και ειδικές εκδόσεις όπως το «Llama Guard 3», σχεδιασμένες για ασφαλή διαχείριση εισροών και εκροών, μειώνοντας τον κίνδυνο παραπλανητικών αποτελεσμάτων.

Επιπλέον, η Meta έχει αρχίσει να πειραματίζεται με πολυμορφικά συστήματα που συνδυάζουν το Llama 3 με δυνατότητες επεξεργασίας εικόνας, βίντεο και ήχου. Αυτά τα πολυμορφικά μοντέλα ανταγωνίζονται τις τρέχουσες εξελίξεις σε εργασίες αναγνώρισης εικόνας, βίντεο και ομιλίας, ενισχύοντας σημαντικά την ευελιξία και τις δυνατότητες εφαρμογής τους πέρα από απλή επεξεργασία κειμένου.

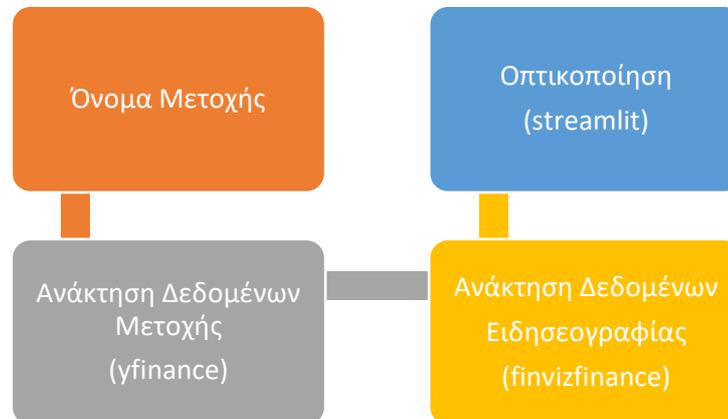
Όσον αφορά τη λειτουργικότητα, το Llama 3 μπορεί να εκτελεί ποικίλες εργασίες όπως παραγωγή συνεκτικού κειμένου (παραγράφων, άρθρων, ιστοριών ή διαλόγων), παροχή ακριβών απαντήσεων σε ερωτήσεις, σύνοψη μεγάλων κειμένων, μετάφραση μεταξύ γλωσσών και διόρθωση προτάσεων ή γραμματικών λαθών. Υποστηρίζει επίσης εφαρμογές ρομπότ συνομιλίας (chatbots), δημιουργία δημιουργικού περιεχομένου (ποιήματα, σενάρια), προγραμματισμό (γεννήτρια και ανάλυση κώδικα) και προσαρμογή σε εξειδικευμένα πεδία όπως νομικά, ιατρικά ή τεχνικά περιεχόμενα.

Σε σύγκριση με το GPT-4, το Llama 3 προσφέρει ιδιαίτερα πλεονεκτήματα σε εργασίες που απαιτούν πολύ μεγάλο περιεχόμενο, ενσωμάτωση μονάδων με εργαλεία, και ενσωμάτωση πολυμορφικών λειτουργιών. Το GPT-4 παραμένει ισχυρό στην πολυγλωσσική κατανόηση, τη σταθερή απόδοση σε ευρεία κλίμακα εργασιών και την ενσωμάτωση μέσω cloud-based υποδομών. Ωστόσο, και τα δύο μοντέλα μοιράζονται κοινά όρια και προκλήσεις, όπως η πιθανότητα παραγωγής μεροληπτικών ή λανθασμένων απαντήσεων και η εξάρτηση από την ποιότητα των προτροπών (prompts).

Συνολικά, το Llama 3 παρουσιάζεται ως ένα ανταγωνιστικό και ευέλικτο LLM που μπορεί να υποκαταστήσει ή να συμπληρώσει το GPT-4 σε πολλές εφαρμογές, ενώ η Meta συνεχίζει την ανάπτυξή του με στόχο την αύξηση της αξιοπιστίας, της πολυμορφικότητας και της προσαρμοστικότητας των μοντέλων της.

## (3.2 Ανάκτηση Ιστορικών Δεδομένων Stock Market)

Η ανάκτηση δεδομένων Stock Market επιτυγχάνεται σύμφωνα με τα βήματα της διαδικασίας όπως παρουσιάζεται στο Σχήμα 3.4.



Σχήμα 3.4: Ανάκτηση δεδομένων

Η εφαρμογή επιτρέπει την ανάκτηση χρηματιστηριακών δεδομένων και ειδήσεων για ένα επιλεγμένο ticker. Ο χρήστης εισάγει το ticker και με το πάτημα ενός κουμπιού η εφαρμογή κατεβάζει τις πρόσφατες ειδήσεις, εμφανίζοντας ημερομηνίες και τίτλους, εξάγει το εύρος ημερομηνιών τους και στη συνέχεια κατεβάζει τις ιστορικές τιμές της μετοχής για το ίδιο διάστημα, υπολογίζοντας την ημερήσια ποσοστιαία μεταβολή. Τα δεδομένα εμφανίζονται σε πίνακες και οι τιμές κλεισίματος οπτικοποιούνται με γραμμικό γράφημα, ενώ η εφαρμογή διαχειρίζεται τυχόν σφάλματα και ενημερώνει τον χρήστη σε πραγματικό χρόνο.

### Περιγραφή Κώδικα

Αρχικά εισάγονται οι απαραίτητες βιβλιοθήκες: streamlit για τη δημιουργία της διαδικτυακής εφαρμογής, yfinance για λήψη ιστορικών χρηματιστηριακών δεδομένων, finvizfinance για λήψη ειδήσεων με βάση το ticker και pandas για χειρισμό και μετατροπή δεδομένων.

```
import streamlit as st
import yfinance as yf
from finvizfinance.quote import finvizfinance
import pandas as pd
```

Στη συνέχεια ορίζεται η συνάρτηση `get_news_data(ticker)` η οποία δημιουργεί αντικείμενο `finvizfinance` για το συγκεκριμένο ticker, ανακτά την ειδησεογραφία (headlines), μετατρέπει τους τίτλους σε πεζά γράμματα, μετατρέπει τη στήλη ημερομηνίας σε μορφή `datetime` και δημιουργεί νέα στήλη που περιέχει μόνο την ημερομηνία χωρίς ώρα.

```
# Function to get news data from Finviz
def get_news_data(ticker):
```

```

stock = finvizfinance(ticker)
news_df = stock.ticker_news()

news_df['Title'] = news_df['Title'].str.lower()
news_df['Date'] = pd.to_datetime(news_df['Date'])
news_df['DateOnly'] = news_df['Date'].dt.date
return news_df

```

Αντιστοίχως, η συνάρτηση `get_stock_data(ticker, start_date, end_date)` χρησιμοποιεί το `yfinance` για να ανακτήσει ιστορικά δεδομένα τιμών από την ημερομηνία `start_date` έως `end_date` και υπολογίζει την ημερήσια ποσοστιαία μεταβολή της τιμής κλεισίματος (`pct_change`).

```

# Function to fetch historical stock data using yfinance
def get_stock_data(ticker, start_date, end_date):
    stock_data = yf.download(ticker, start=start_date, end=end_date)
    stock_data['Pct_Change'] = stock_data['Close'].pct_change() * 100
    return stock_data

```

Η διεπαφή της εφαρμογής με το `streamlit` περιλαμβάνει τον τίτλο της εφαρμογής "📰 Stock News Viewer" και ένα πλαίσιο εισαγωγής κειμένου για το `ticker` με προεπιλεγμένη τιμή, για παράδειγμα "SBUX". Όταν ο χρήστης πατάει το κουμπί για λήψη δεδομένων, εμφανίζεται μήνυμα πληροφόρησης ότι γίνεται λήψη ειδήσεων και καλείται η συνάρτηση `get_news_data` για απόκτηση των ειδήσεων. Αν δεν βρεθούν ειδήσεις, εμφανίζεται προειδοποίηση, ενώ αν υπάρχουν, εμφανίζεται πίνακας με τις ημερομηνίες και τους τίτλους των ειδήσεων. Από τις ειδήσεις εξάγεται το εύρος ημερομηνιών και εμφανίζεται μήνυμα για λήψη ιστορικών δεδομένων τιμών, τα οποία ανακτώνται με την `get_stock_data`. Τα δεδομένα τιμών εμφανίζονται σε πίνακα με τιμές κλεισίματος και ημερήσιες ποσοστιαίες μεταβολές, ενώ οι τιμές κλεισίματος απεικονίζονται σε γραμμικό γράφημα. Η εφαρμογή περιλαμβάνει επίσης διαχείριση σφαλμάτων για ενημέρωση του χρήστη σε περίπτωση προβλημάτων.

```

# Streamlit UI section starts here
st.title("📰 Stock News Viewer")

# Streamlit User Interface section starts here
st.title("📰 Stock News Viewer")

# Input box for user to enter stock ticker
ticker = st.text_input("Enter stock ticker (e.g., SBUX)", "SBUX")

# Button to trigger data fetching
if st.button("Get News and Stock Data"):
    try:
        st.info(f"Fetching news for {ticker}...")

```

```

news_df = get_news_data(ticker)

if news_df.empty:
    st.warning("No news data found.")
else:
    st.subheader("📰 News Headlines")
    st.dataframe(news_df[["DateOnly", 'Title']])

    # Determine date range based on news
    start_date = news_df['DateOnly'].min().strftime('%Y-%m-%d')
    end_date = news_df['DateOnly'].max().strftime('%Y-%m-%d')

    st.info(f"Fetching stock data from {start_date} to {end_date}...")
    stock_data = get_stock_data(ticker, start_date, end_date)

    st.subheader("📊 Stock Prices and Daily Change")
    st.dataframe(stock_data[["Close", 'Pct_Change']].dropna())

except Exception as e:
    st.error(f"Error: {e}")

```

### Αποτελέσματα κώδικα

Η εφαρμογή finvizfinance επιτρέπει τη λήψη μέχρι περίπου 100 ειδήσεων ανά ticker, ανεξαρτήτως ημερομηνίας. Αυτό σημαίνει ότι αν σε μία ημέρα εκδοθούν πολλές ειδήσεις, το σύνολο τους μπορεί να καλύπτει λιγότερο από ένα μήνα. Για να συλλεχθεί ένας σημαντικός όγκος δεδομένων, απαιτείται συνεχής ανάκτηση και αποθήκευση των ειδήσεων σε καθημερινή βάση, ώστε στη συνέχεια να μπορούν να συμπυκνωθούν και να αναλυθούν. Αυτός ο περιορισμός πρέπει να λαμβάνεται υπόψη κατά την ανάπτυξη συστημάτων ανάλυσης ειδήσεων και ιστορικών δεδομένων.

Στον Πίνακα 3.3 παρουσιάζεται μέρος των αποτελεσμάτων για την ειδησεογραφία σχετικά με την μετοχή SBUX. Συνολικά επιστρέφονται 100 ειδήσεις από τις πιο πρόσφατες και πίσω.

**Πίνακας 3.3:** Παράδειγμα ανάκτησης ειδήσεων για SBUX (2025-10-23 έως 2025-10-30)

Ημερομηνία	Είδηση
30/10/2025	analyst explains why he's bullish on starbucks (sbux), says ceo brian niccol's plan is working
30/10/2025	earnings live: amazon stock soars, apple ceo predicts 'best-ever' holiday quarter, reddit and coinbase rise
30/10/2025	earnings live: amazon stock soars on aws growth, apple wavers, reddit and coinbase move higher
30/10/2025	starbucks is carrying out a major revamp. there are signs it's working.
30/10/2025	chipotle stock is crashing 15% today. should you buy the dip?
30/10/2025	starbucks sales grow, chipotle's consumer woes: earnings takeaways

Ημερομηνία	Είδηση
30/10/2025	starbucks q4 earnings miss estimates, revenues increase y/y
30/10/2025	behind the rise, fall and reinvention of unions
30/10/2025	starbucks ceo on turnaround plan: i believe we are ahead of schedule
30/10/2025	starbucks plans 1,000 store renovations by end of 2026 and tests out new budget prototype
30/10/2025	'back to starbucks' turnaround plan is working, ceo niccol says
30/10/2025	baristas are embracing changes at starbucks, ceo says
30/10/2025	chipotle stock dives on consumer woes. why starbucks was unfazed.
30/10/2025	starbucks terrible quarter
30/10/2025	analyst report: starbucks corporation
30/10/2025	starbucks ceo: believe we are ahead of schedule on turnaround plan
30/10/2025	starbucks stock drops. earnings offer a glimpse at companys turnaround plans.
30/10/2025	earnings live: meta stock tumbles, microsoft slides, and alphabet jumps as big tech earnings pour in
30/10/2025	starbucks says its turnaround is starting to work
30/10/2025	markets wanted a boost from earnings, fed and china talks. how powell disappointed.
30/10/2025	starbucks returns to samestore sales growth, led by international markets
30/10/2025	starbucks stock: analyst estimates & ratings
30/10/2025	target's next ceo finds a bold solution to alarming sales declines

Τα δεδομένα του χρηματιστηρίου (stock market) είναι διαθέσιμα μόνο για εργάσιμες ημέρες, δηλαδή ημέρες κατά τις οποίες τα χρηματιστήρια είναι ανοικτά, ένα σετ την ημέρα. Αυτό σημαίνει ότι δεν υπάρχουν τιμές για σαββατοκύριακα ή αργίες, κάτι που πρέπει να λαμβάνεται υπόψη κατά την ανάλυση των ιστορικών δεδομένων.

Συνδυάζοντας τα δεδομένα ειδησεογραφίας με τα δεδομένα του χρηματιστηρίου, στην περίπτωση αυτή ανακτήθηκαν δεδομένα για 6 ημέρες, καθώς οι ειδήσεις του finvizfinance περιορίζονται σε περίπου 100 ανά ticker και συχνά καλύπτουν λιγότερο από ένα μήνα. Αντίθετα, τα ιστορικά δεδομένα τιμών από το yfinance δεν έχουν όριο στον αριθμό ημερών που μπορούν να ανακτηθούν, αρκεί να είναι διαθέσιμα στο χρηματιστήριο.

**Πίνακας 3.4:** Αποτέλεσμα ανάκτησης δεδομένων για SBUX (Starbucks)

Ημερομηνία	Τιμή Κλεισίματος	Ημερήσια ποσοστιαία μεταβολή
2025-10-23	85.44	-0.07
2025-10-24	86.09	0.76
2025-10-27	87.22	1.31
2025-10-28	85.43	-2.05
2025-10-29	84.17	-1.47
2025-10-30	83.15	-1.21

## (3.3 Ανάλυση Συναισθήματος)

---

### (3.3.a Ανάλυση Συναισθήματος στο Χρηματιστήριο)

---

Η ανάλυση συναισθήματος (sentiment analysis) αποτελεί μια τεχνική της τεχνητής νοημοσύνης και πιο συγκεκριμένα της επεξεργασίας φυσικής γλώσσας (NLP), η οποία επιτρέπει την αξιολόγηση κειμένων, όπως ειδήσεων, άρθρων ή αναρτήσεις στα μέσα κοινωνικής δικτύωσης, για την αναγνώριση θετικού, αρνητικού ή ουδέτερου συναισθήματος [6],[7]. Στο χρηματιστήριο, η ανάλυση αυτή μπορεί να αποκαλύψει την ψυχολογία των επενδυτών και την αντίληψη της αγοράς για μια συγκεκριμένη μετοχή ή κλάδο, επηρεάζοντας την προσφορά, τη ζήτηση και τελικά την τιμή των μετοχών [6],[7]. Η σημασία της ανάλυσης συναισθήματος έγκειται στο γεγονός ότι οι χρηματιστηριακές τιμές δεν καθορίζονται μόνο από θεμελιώδη οικονομικά μεγέθη, αλλά και από ψυχολογικούς παράγοντες όπως φόβος, αισιοδοξία ή φήμη [6].

Η διαδικασία ανάλυσης συναισθήματος περιλαμβάνει αρκετά στάδια επεξεργασίας κειμένου, όπως: καθαρισμό δεδομένων, κατάτμηση σε λεκτικές μονάδες (tokenization), αφαίρεση κοινών λέξεων (stopwords), λεμματοποίηση ή αποκοπή καταλήξεων (stemming) και αναγνώριση ονομάτων εταιριών ή προσώπων (Named Entity Recognition), προκειμένου να εντοπιστούν τα σχετικά στοιχεία μέσα στο κείμενο [6],[7],[14]. Στη συνέχεια, εφαρμόζονται διάφορες μέθοδοι ανάλυσης. Οι κυριότερες προσεγγίσεις περιλαμβάνουν ανάλυση με βάση λεξικό, όπου χρησιμοποιούνται προκαθορισμένες λίστες λέξεων με συναισθηματική χροιά, μεθόδους μηχανικής μάθησης όπως Random Forests [8], Bayesian Classifiers [9], Extremely Randomized Trees [10], Support Vector Machines [11] και Gradient Boosting [12], καθώς και μεθόδους βαθιάς μάθησης που χρησιμοποιούν μοντέλα μεγάλης κλίμακας όπως BERT ή FinBERT [1], [2], [3], [4], .

Η οπτικοποίηση των αποτελεσμάτων γίνεται με συγκέντρωση σκορ ανά μετοχή ή ημερομηνία, δημιουργία γραφημάτων ή θερμικών χαρτών συναισθήματος (heatmaps), διευκολύνοντας την ανάλυση των τάσεων της αγοράς [5]. Ωστόσο, η μέθοδος παρουσιάζει περιορισμούς, καθώς η ερμηνεία ειρωνείας και σαρκασμού παραμένει δύσκολη, ειδικά σε δεδομένα από μέσα κοινωνικής δικτύωσης, ενώ η ύπαρξη θορύβου ή παραπλανητικών πληροφοριών μπορεί να επηρεάσει την ακρίβεια των αποτελεσμάτων. Επιπλέον, η ανάλυση συναισθήματος δεν αντικαθιστά την θεμελιώδη ανάλυση και είναι πιο κατάλληλη για βραχυπρόθεσμες επενδυτικές κινήσεις, με έμφαση συνήθως σε μεγάλες και δημοφιλείς μετοχές [6], [7].

Η μέθοδος ανάλυσης συναισθήματος παρουσιάζει ορισμένους περιορισμούς. Η ερμηνεία ειρωνείας και σαρκασμού παραμένει δύσκολη, ιδιαίτερα σε κείμενα από μέσα κοινωνικής δικτύωσης, όπου η γλώσσα είναι λιγότερο τυπική. Επιπλέον, η ύπαρξη θορύβου λόγω άχρηστων ή παραπλανητικών δεδομένων μπορεί να επηρεάσει την ακρίβεια και την αξιοπιστία των αποτελεσμάτων. Η ανάλυση συναισθήματος δεν υποκαθιστά την θεμελιώδη ανάλυση των εταιρικών δεδομένων και είναι περισσότερο κατάλληλη για βραχυπρόθεσμες επενδυτικές αποφάσεις. Επίσης, η εφαρμογή της επικεντρώνεται κυρίως σε μεγάλες και δημοφιλείς μετοχές, με αποτέλεσμα να παραβλέπονται μικρότερες εταιρίες.

Παρά τους περιορισμούς αυτούς, η ανάλυση συναισθήματος προσφέρει έναν σύγχρονο τρόπο μέτρησης της αντίληψης και της ψυχολογίας της αγοράς, ειδικά σε περιόδους έντονης ειδησεογραφίας ή αυξημένης αβεβαιότητας. Όταν συνδυαστεί με θεμελιώδη και τεχνική ανάλυση, μπορεί να παρέχει σημαντικό ανταγωνιστικό πλεονέκτημα στους επενδυτές.

### (3.3.β Κατηγοριοποίηση Συναισθήματος για Stock Market)

Σε συνέχεια της εφαρμογής που παρουσιάζεται στην παράγραφο 3.2, προστίθεται η δυνατότητα ανάλυσης συναισθήματος με τη βοήθεια του μοντέλου Llama 3 της πλατφόρμας Ollama. Το μοντέλο αυτό, ως Large Language Model (LLM), επιτρέπει την αξιολόγηση των ειδήσεων και των χρηματιστηριακών δεδομένων για τον εντοπισμό θετικών, αρνητικών ή ουδέτερων συναισθημάτων, συμβάλλοντας στην καλύτερη κατανόηση της ψυχολογίας της αγοράς και των πιθανών βραχυπρόθεσμων κινήσεων των τιμών.

#### Περιγραφή Κώδικα

Η εφαρμογή αρχίζει με τον ορισμό μιας προτροπής (prompt) που ζητά από το μοντέλο Llama 3 να ταξινομήσει το συναίσθημα ενός τίτλου ειδήσεων ως θετικό, αρνητικό ή ουδέτερο και να δώσει μια σύντομη αιτιολόγηση για την επιλογή του. Το prompt στέλνεται στο τοπικό API του Ollama μέσω ενός αιτήματος HTTP POST, με χρονικό όριο για την απάντηση ώστε να αποφευχθεί η ατέρμονη αναμονή. Αν το API επιστρέψει επιτυχώς την απάντηση (HTTP 200), το αποτέλεσμα ανακτάται, καθαρίζεται και μετατρέπεται σε πεζά γράμματα. Η ταξινόμηση πραγματοποιείται με έλεγχο λέξεων-κλειδιών, όπου αν η απάντηση περιέχει τη λέξη "positive" επιστρέφεται η κατηγορία "Positive", αν περιέχει "negative" επιστρέφεται "Negative", ενώ σε κάθε άλλη περίπτωση επιστρέφεται "Neutral". Σε περίπτωση καθυστέρησης ή άλλου σφάλματος, επιστρέφεται το αντίστοιχο μήνυμα "Timeout" ή "Error".

```
prompt = f"Classify the sentiment of this news headline as Positive, Negative, or Neutral. Provide a brief justification for the classification:\n\n\"{text}\"\\n\nSentiment and Justification:"
response = requests.post(
    'http://localhost:11434/api/generate',
    json={"model": model, "prompt": prompt, "stream": False},
    timeout=timeout # Timeout to avoid infinite wait
)
if response.status_code == 200:
    result = response.json()["response"].strip().lower()
    if 'positive' in result:
        return "Positive"
    elif 'negative' in result:
        return "Negative"
    else:
        return "Neutral"
except requests.exceptions.Timeout:
    return "Timeout"
except Exception:
    return "Error"
```

Ως περιβάλλον χρήστη της εφαρμογής, χρησιμοποιείται το streamlit με τίτλο "Stock News Sentiment Classifier" και πλαίσιο εισαγωγής για τον χρήστη ώστε να εισάγει το σύμβολο της μετοχής. Όταν ο χρήστης πατά το κουμπί "Get News and Sentiment Classification", ανακτώνται οι τελευταίες ειδήσεις για το συγκεκριμένο ticker μέσω της συνάρτησης `get_news_data(ticker)`. Αν δεν υπάρχουν ειδήσεις εμφανίζεται προειδοποίηση. Σε περίπτωση ύπαρξης ειδήσεων, περιορίζονται στις πρώτες

δέκα για λόγους ταχύτητας και κάθε τίτλος υποβάλλεται σε ταξινόμηση συναισθήματος μέσω της συνάρτησης `classify_sentiment_with_justification()`, η οποία επιστρέφει τόσο την κατηγορία συναισθήματος όσο και την αιτιολόγηση. Κατά τη διάρκεια της διαδικασίας εμφανίζεται γραφικό προόδου (progress bar).

Τα αποτελέσματα αποθηκεύονται σε ένα πίνακα δεδομένων (DataFrame) που περιλαμβάνει την ημερομηνία, τον τίτλο, την κατηγορία συναισθήματος με χρωματική κωδικοποίηση ή emoji, και την αιτιολόγηση για κάθε τίτλο. Στο τελικό περιβάλλον εμφανίζεται ο πίνακας με τις ειδήσεις και τα συναισθήματα καθώς και η πλήρης αιτιολόγηση ανά τίτλο, δίνοντας στον χρήστη σαφή εικόνα για το συναίσθημα της αγοράς σε σχέση με τις ειδήσεις της συγκεκριμένης μετοχής.

```
st.title("📰 Stock News Sentiment Classifier")
# Input field for the stock ticker symbol (e.g., AAPL, TSLA, SBUX)
ticker = st.text_input("Enter stock ticker (e.g., SBUX)", "SBUX")
st.title("📰 Stock News Sentiment Classifier")

# Input field for the stock ticker symbol (e.g., AAPL, TSLA, SBUX)
ticker = st.text_input("Enter stock ticker (e.g., SBUX)", "SBUX")

# When the button is clicked, the app fetches and displays data
if st.button("Get News and Sentiment Classification"):
    try:
        st.info(f"Fetching news for {ticker}...")

        # Fetch recent news headlines
        news_df = get_news_data(ticker)

        if news_df.empty:
            st.warning("No news data found.")
        else:
            # Limit to first 10 headlines to speed up classification
            st.info("Classifying sentiment using Ollama (first 10 headlines)...")
            news_df = news_df.head(10)

            # Display progress bar during sentiment classification
            progress = st.progress(0)
            sentiments = []
            justifications = []

            # Loop through each headline and classify sentiment
            for i, title in enumerate(news_df['Title']):
                sentiment, justification = classify_sentiment_with_justification(title)
                sentiments.append(sentiment)
                justifications.append(justification)
                progress.progress((i + 1) / len(news_df)) # Update progress bar
```

```

# Add sentiment and justification results to the DataFrame
news_df['Sentiment'] = sentiments
news_df['Justification'] = justifications

# Add emojis/colors for visual feedback
def color_sentiment(sentiment):
    if sentiment == "Positive":
        return "🟢 Positive"
    elif sentiment == "Negative":
        return "🔴 Negative"
    elif sentiment == "Neutral":
        return "⚪ Neutral"
    else:
        return f"⚠️ {sentiment}"

# Apply color-coded labels
news_df['Sentiment_Colored'] = news_df['Sentiment'].apply(color_sentiment)

# Display the final news DataFrame with colored sentiment
st.subheader("📰 News Headlines with Sentiment")
st.dataframe(news_df[['DateOnly', 'Title', 'Sentiment_Colored']].rename(
    columns={'DateOnly': 'Date', 'Sentiment_Colored': 'Sentiment'}
))

# Display sentiment answers and justification
st.subheader("💬 Sentiment Classification and Justification:")
for idx, row in news_df.iterrows():
    st.markdown(f"Headline: {row['Title']}")
    st.markdown(f"Sentiment: {row['Sentiment']}")
    st.markdown(f"Justification: {row['Justification']}")
    st.markdown("---")

except Exception as e:
    st.error(f"Error: {e}") # Catch any exception and show error message

```

Ο Πίνακας 3.5 παρουσιάζει ένα παράδειγμα κατηγοριοποίησης συναισθήματος για τη μετοχή SBUX (Starbucks) σε 10 ειδήσεις του Οκτωβρίου 2025. Ο πίνακας εμφανίζει την ημερομηνία, τον τίτλο της είδησης και την αντίστοιχη κατηγορία συναισθήματος που αποδόθηκε από το μοντέλο Llama 3 μέσω της συνάρτησης ταξινόμησης συναισθήματος.

Πίνακας 3.5: Παράδειγμα κατηγοριοποίησης συναισθήματος για SBUX (Starbucks) σε 10 ειδήσεις Οκτώβριος 2025

Ημερομηνία	Είδηση	Συναίσθημα
31/10/2025	starbucks ceo brian niccol rules out a value menu as chain revamps	Neutral
31/10/2025	earnings live: amazon stock soars, apple pares gains with earnings from ai players to come	Neutral
31/10/2025	sbux q4 earnings sees first global comp growth in seven quarters	Positive
31/10/2025	what's brewing for starbucks after mixed q4 earnings? etfs in focus	Neutral
31/10/2025	earnings live: amazon stock soars, apple ceo predicts 'best-ever' quarter; exxon, chevron post strong earnings	Positive
31/10/2025	chipotle struggles as customers skip the guac	Negative
30/10/2025	analyst explains why he's bullish on starbucks (sbux), says ceo brian niccol's plan is working	Positive
30/10/2025	earnings live: amazon stock soars, apple ceo predicts 'best-ever' holiday quarter, reddit and coinbase rise	Positive
30/10/2025	earnings live: amazon stock soars on aws growth, apple wavers, reddit and coinbase move higher	Positive
30/10/2025	starbucks is carrying out a major revamp. there are signs it's working.	Positive

Παρακάτω παρουσιάζεται η αιτιολόγηση που παρέχει το Llama 3 για τρεις τύπους συναισθήματος: ένα ουδέτερο, ένα θετικό και ένα αρνητικό. Αυτή η αιτιολόγηση εξηγεί γιατί κάθε τίτλος χαρακτηρίστηκε με συγκεκριμένο τρόπο, δίνοντας στον αναγνώστη σαφή εικόνα της λογικής που ακολουθεί το μοντέλο για την εκτίμηση της ψυχολογίας της αγοράς.

#### Ουδέτερο (Neutral)

**Είδηση:** «Τι «βράζει» για τη Starbucks μετά τα μεικτά αποτελέσματα του τέταρτου τριμήνου; Επενδυτικά Ταμεία (ETFs) στο επίκεντρο»

**Ανάλυση συναισθήματος:** Ουδέτερο

**Δικαιολόγηση:** Ο τίτλος απλώς αναφέρει τα αποτελέσματα του τέταρτου τριμήνου της Starbucks και αναφέρει τα ETFs ως σχετικό θέμα, χωρίς να εκφράζει σαφές θετικό ή αρνητικό συναίσθημα. Πρόκειται περισσότερο για ενημερωτικό τίτλο παρά για άρθρο γνώμης, γεγονός που αιτιολογεί την ταξινόμησή του ως ουδέτερο (Neutral).

**Headline:** what's brewing for starbucks after mixed g4 earnings? etfs in focus

**Sentiment:** Neutral

**Justification:** Justification: The headline is simply reporting on Starbucks' Q4 earnings and mentioning ETFs (Exchange-Traded Funds) as a related topic, without expressing any clear positive or negative sentiment. It's more of an informative title than an opinion piece, which justifies its classification as Neutral.

### Θετικό (Positive)

**Είδηση:** «Ο αναλυτής εξηγεί γιατί είναι αισιόδοξος για τη Starbucks (SBUX) και λέει ότι το σχέδιο του CEO Brian Niccol αποδίδει»

**Ανάλυση συναισθήματος:** Θετικό

**Δικαιολόγηση:** Ο τίτλος αναφέρει ότι ο αναλυτής είναι αισιόδοξος (bullish) για τη μετοχή της Starbucks, κάτι που υποδηλώνει θετικό συναίσθημα. Επιπλέον, το γεγονός ότι ο αναλυτής επαινεί το σχέδιο του CEO Brian Niccol ως "λειτουργεί" δείχνει θετικές προοπτικές για την εταιρεία, γεγονός που καθιστά το συναίσθημα του τίτλου συνολικά θετικό.

**Headline:** analyst explains why he's bullish on starbucks (sbux), says ceo brian niccol's plan is working

**Sentiment:** Positive

**Justification:** The headline mentions being "bullish" on Starbucks, which is a positive sentiment. Additionally, the fact that an analyst is praising CEO Brian Niccol's plan as "working" suggests a positive outlook for the company, indicating that the sentiment of this headline is overall positive.

### Αρνητικό (Negative)

**Είδηση:** «Η Chipotle αντιμετωπίζει δυσκολίες καθώς οι πελάτες παραλείπουν το γουακαμόλε»

**Ανάλυση συναισθήματος:** Αρνητικό

**Δικαιολόγηση:** Η χρήση της λέξης "δυσκολεύεται" υποδηλώνει ότι η Chipotle αντιμετωπίζει δυσκολίες, κάτι που έχει αρνητική χροιά. Επιπλέον, η φράση "οι πελάτες παραλείπουν το γουακαμόλε" δείχνει ότι οι πελάτες δεν αγοράζουν το χαρακτηριστικό προϊόν της εταιρείας (guacamole), κάτι που πιθανώς επιδεινώνει τα προβλήματά της. Συνολικά, ο τόνος του τίτλου είναι επικριτικός και υποδηλώνει ότι κάτι δεν πάει καλά για την Chipotle, γεγονός που το καθιστά αρνητικό (Negative).

**Headline:** chipotle struggles as customers skip the guac

**Sentiment:** Negative

**Justification:** The use of the word "struggles" implies that Chipotle is experiencing difficulties, which has a negative connotation. Additionally, the phrase "customers skip the guac" suggests that customers are not choosing to buy Chipotle's signature item (guacamole), which is likely a significant contributor to their struggles. Overall, the tone of the headline is critical and suggests that something is going wrong for Chipotle, which makes it a negative sentiment.

**Σημείωση ως προς τη σχετική είδηση.** Αναρωτιέται κανείς γιατί το finvizfinance έδωσε τη συγκεκριμένη είδηση σε σχέση με την Starbucks. Το πιο πιθανό είναι ότι αναφέρεται στον ίδιο κλάδο καφετέριες /fast food και αφορά τάσεις στο σχετικό τομέα. Η ανάλυση συναισθήματος παρόλα αυτά αξιολογείται σωστή.

### (3.3.γ Οπτικοποίηση Συναισθήματος)

#### *heatmap*

Ένας θερμοχάρτης (heatmap) αποτελεί ένα γραφικό εργαλείο απεικόνισης δεδομένων, όπου οι τιμές αναπαρίστανται μέσω χρωμάτων, επιτρέποντας την οπτική αναγνώριση προτύπων, διαφορών και ακραίων τιμών [5]. Συνήθως χρησιμοποιείται για την παρουσίαση πινάκων, όπως πίνακες συσχέτισεων, συχνοτήτων ή αποδόσεων, ώστε να διευκολυνθεί η κατανόηση των δεδομένων. Σε έναν θερμοχάρτη, κάθε κελί αντιστοιχεί σε μία τιμή, ενώ το χρώμα του δείχνει το μέγεθός της, πιο φωτεινά ή έντονα χρώματα αντιπροσωπεύουν μεγαλύτερες τιμές και πιο σκούρα ή ανοιχτά χρώματα μικρότερες. Συχνά εφαρμόζεται χρωματική κλίμακα (gradient) για να αναδειχθούν οι θετικές και αρνητικές τιμές.

Οι θερμοχάρτες χρησιμοποιούνται ευρέως για την οπτικοποίηση σύνθετων δεδομένων, με στόχο την αναγνώριση προτύπων, συσχέτισεων και αποκλίσεων. Αποδεικνύονται ιδιαίτερα χρήσιμοι στην ανάλυση συσχέτισεων μεταξύ μεταβλητών, στην παρακολούθηση της απόδοσης σε διαφορετικές χρονικές περιόδους και στην απεικόνιση συχνοτήτων ή πυκνοτήτων. Επιπλέον, βρίσκουν εφαρμογή σε χρηματοοικονομικά, βιοπληροφορική, επιστήμη δεδομένων και παρακολούθηση συστημάτων, διευκολύνοντας την ερμηνεία πολύπλοκων πινάκων μέσω της χρήσης χρωματικών κλιμάκων. Στην ανάπτυξη κώδικα σε Python, η βιβλιοθήκη seaborn προσφέρει έτοιμη συνάρτηση για τη δημιουργία θερμοχάρτων (seaborn.heatmap).

Ο κώδικας δημιουργεί έναν θερμοχάρτη (heatmap) που απεικονίζει τη μέση βαθμολογία συναισθήματος των ειδήσεων ανά ημέρα. Αρχικά δημιουργείται ένας πίνακας σύνοψης με ημερομηνίες ως δείκτες και τις μέσες τιμές του σκορ συναισθήματος. Στη συνέχεια, χρησιμοποιώντας τη βιβλιοθήκη seaborn, σχεδιάζεται ο heatmap με χρωματική κλίμακα “ψυχρών-θερμών χρωμάτων” για να αναδεικνύονται οι θετικές και αρνητικές τιμές. Το γράφημα εμφανίζεται στην εφαρμογή streamlit με τη συνάρτηση st.pyplot(fig), επιτρέποντας στον χρήστη να βλέπει οπτικά τις τάσεις συναισθήματος των ειδήσεων ανά ημέρα.

```
# Display sentiment heatmap
st.subheader("🔥 Sentiment Heatmap")

# Create figure and axis objects for plotting
fig, ax = plt.subplots(figsize=(10, 6))

# Prepare sentiment pivot table
sentiment_pivot = news_df.pivot_table(values='Sentiment_Score', index='DateOnly',
aggfunc='mean')

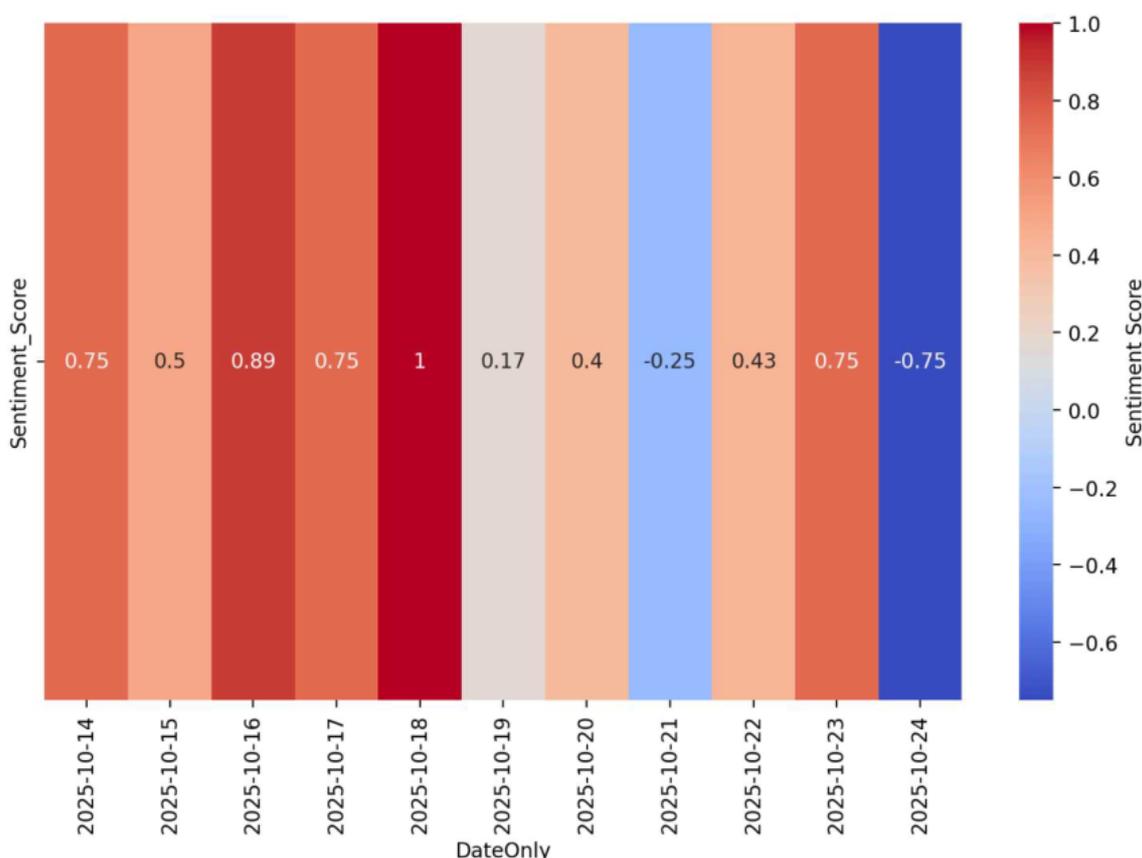
# Plot the heatmap using seaborn
sns.heatmap(sentiment_pivot.T, annot=True, cmap="coolwarm", cbar_kws={'label': 'Sentiment
Score'}, ax=ax)

# Display the heatmap in Streamlit
st.pyplot(fig) # Provide the figure explicitly
```

Σε συνέχεια της εφαρμογής της παραγράφου 3.3.β, και για 100 ειδήσεις για SBUX (το μέγιστο δυνατό), από finvizfinance γίνεται ανάλυση συναισθήματος από Llama 3 για κάθε είδηση.

Για οπτικοποίηση της συνολικής διάθεσης της αγοράς, οι κατηγορίες συναισθήματος μετατρέπονται σε αριθμητικές τιμές (“Θετικό”=1, “Ουδέτερο”=0, “Αρνητικό”=-1) και υπολογίζεται η μέση τιμή ανά ημέρα.

Οι ημερήσιες τιμές που προκύπτουν απεικονίζονται σε θερμοχάρτη (heatmap) με χρωματική κλίμακα, όπου τα θερμά χρώματα δείχνουν θετικές τάσεις και τα ψυχρά αρνητικές, διευκολύνοντας την ταχεία αναγνώριση προτύπων, αποκλίσεων και περιόδων αβεβαιότητας.



Σχήμα 3.5: Παράδειγμα θερμοχάρτη για SBUX τελευταίες ειδήσεις (Οκτ 2025)

Από τις 14 έως τις 18 Οκτωβρίου παρουσιάζει σταθερά θετικό συναίσθημα, υποδηλώνοντας αυξημένη εμπιστοσύνη, θετικές ειδήσεις ή ισχυρή αγοραστική διάθεση. Στις 19–21 Οκτωβρίου εμφανίζεται σαφής επιδείνωση, με κορύφωση την αρνητική τιμή της 21ης, κάτι που συνήθως αντιστοιχεί σε αρνητικές ειδήσεις, πτώση προσδοκιών ή αυξημένη αβεβαιότητα. Η αγορά φαίνεται να αντιδρά βραχυπρόθεσμα θετικά στις 22–23 Οκτωβρίου, όμως η κατακόρυφη αρνητική μεταστροφή της 24ης Οκτωβρίου δηλώνει πιθανή ανακοίνωση αρνητικών αποτελεσμάτων, γεγονός υψηλού κινδύνου ή έντονη διόρθωση. Συνολικά, το μοτίβο δείχνει μια μετοχή με κυκλικό και ευαίσθητο συναίσθημα, όπου τα θετικά κύματα είναι ισχυρά αλλά ακολουθούνται από απότομα αρνητικά γυρίσματα, κάτι που συνήθως χαρακτηρίζει τίτλους υψηλής μεταβλητότητας ή έντονης ειδησεογραφικής επίδρασης.

### (3.4 Πρόβλεψη Απόδοσης Μετοχής)

---

Η πρόβλεψη των χρηματιστηριακών τιμών αποτελεί ένα από τα πιο σύνθετα και απαιτητικά προβλήματα της επιστήμης δεδομένων, καθώς οι αγορές επηρεάζονται από πλήθος δυναμικών παραγόντων, όπως οικονομικοί δείκτες, ειδησεογραφικές εξελίξεις, επενδυτικό συναίσθημα και απρόβλεπτα γεωπολιτικά γεγονότα. Η εφαρμογή μεγάλων γλωσσικών μοντέλων στις χρηματοοικονομικές προβλέψεις έχει ήδη δείξει ότι μπορεί να ενισχύσει την ανάλυση επενδυτικών χαρτοφυλακίων, τις αξιολογήσεις μετοχών και την πρόβλεψη τάσεων της αγοράς [6], [7].

Ο συνδυασμός κλασικών τεχνικών μηχανικής μάθησης με σύγχρονες μεθοδολογίες LLM δημιουργεί ένα ισχυρό πλαίσιο για την πρόβλεψη βραχυπρόθεσμων μεταβολών της τιμής μετοχών. Η προσέγγιση αυτή αξιοποιεί τόσο την ιστορική στατιστική συμπεριφορά της αγοράς όσο και την πληροφορία που κρύβεται σε μεγάλα σύνολα ειδησεογραφικού περιεχομένου, προσφέροντας μια πιο ολοκληρωμένη και ρεαλιστική εικόνα της δυναμικής των χρηματιστηριακών αγορών.

Με αξιοποίηση των εργαλείων και μεθόδων που παρουσιάστηκαν στις προηγούμενες ενότητες, αναπτύχθηκε μια διαδραστική εφαρμογή με στόχο την πρόβλεψη της ημερήσιας απόδοσης μετοχών. Η εφαρμογή φορτώνει και επεξεργάζεται δεδομένα ειδήσεων καθώς και ιστορικά δεδομένα τιμών της μετοχής για το ίδιο χρονικό διάστημα, επιτρέποντας τον συσχετισμό εξωγενών πληροφοριών με χρηματιστηριακές μεταβολές. Για την ανάλυση του συναισθήματος των τίτλων ειδήσεων αξιοποιείται το τοπικό API του Ollama σε συνδυασμό με το γλωσσικό μοντέλο Llama 3, παρέχοντας μια αποδοτική και ιδιωτική διαδικασία εξαγωγής χαρακτηριστικών. Στη συνέχεια, τα δεδομένα ειδήσεων ενοποιούνται με τα ημερήσια δεδομένα της μετοχής και χρησιμοποιούνται για την εκπαίδευση μοντέλων μηχανικής μάθησης, τα οποία προβλέπουν την ποσοστιαία μεταβολή της τιμής. Τέλος, η εφαρμογή παρουσιάζει μετρικές απόδοσης και διαγράμματα που συγκρίνουν τις προβλέψεις με τις πραγματικές τιμές, προσφέροντας μια ολοκληρωμένη εικόνα της αποτελεσματικότητας του συστήματος.

#### (3.4.α Προετοιμασία Δεδομένων)

---

Δεδομένου ότι η πρόβλεψη χρηματιστηριακής απόδοσης απαιτεί μεγάλο όγκο ιστορικών δεδομένων, η χρήση της πλατφόρμας Finviz αποδείχθηκε ανεπαρκής, καθώς επιβάλλει περιορισμό έως 100 ειδήσεων ανά ανάκτηση. Το όριο αυτό είναι ιδιαίτερα περιοριστικό, αφού πολλές φορές ο ημερήσιος όγκος ειδήσεων για μια μετοχή μπορεί να υπερβαίνει αυτόν τον αριθμό, με αποτέλεσμα η μέγιστη περίοδος κάλυψης να περιορίζεται περίπου σε έναν μήνα. Για τον λόγο αυτό, η αναζήτηση μεγαλύτερων συνόλων δεδομένων οδήγησε στην αξιοποίηση ιστορικού αρχείου ειδήσεων που διατέθηκε μέσω της ακόλουθης πηγής της κοινότητας αλγοριθμικών συναλλαγών: [https://www.reddit.com/r/algorithmictrading/comments/hl3x5j/i\\_compiled\\_data\\_for\\_4m\\_stock\\_news\\_articles\\_for/](https://www.reddit.com/r/algorithmictrading/comments/hl3x5j/i_compiled_data_for_4m_stock_news_articles_for/). Χρησιμοποιήθηκαν δεδομένα καλύπτοντας τρία χρόνια, από το 2018 έως το 2020, όσο πιο πολλά σημεία και πιο πρόσφατα.

Ο Πίνακας 3.6 παρουσιάζει ένα ενδεικτικό σύνολο ημερήσιων δεδομένων για τη μετοχή της Starbucks (SBUX), όπως ανακτήθηκαν μέσω της βιβλιοθήκης yfinance. Για κάθε ημερομηνία περιλαμβάνονται οι βασικές χρηματιστηριακές μεταβλητές (τιμή έναρξης, υψηλή, χαμηλή, τιμή κλεισίματος και όγκος συναλλαγών) καθώς και δύο υπολογιζόμενες μεταβλητές: η ποσοστιαία ημερήσια μεταβολή της τιμής και η ποσοστιαία μεταβολή της προηγούμενης ημέρας. Τα δεδομένα αυτά χρησιμοποιούνται στη συνέχεια ως είσοδος για την εκπαίδευση των μοντέλων μηχανικής μάθησης, επιτρέποντας την αποτύπωση της δυναμικής της χρονοσειράς και της βραχυπρόθεσμης μεταβλητότητας της αγοράς.

**Πίνακας 3.6:** Δεδομένα από το yfinance για SBUX (ενδεικτικό δείγμα του DataFrame)

Ημερομηνία	Τιμή Κλεισίματος	Υψηλή Τιμή	Χαμηλή Τιμή	Τιμή Έναρξης	Όγκος	Ποσοστιαία Ημερήσια Μεταβολή Τιμής	Προηγούμενη Ποσοστιαία Ημερήσια Μεταβολή Τιμής
24/01/2018	51.46	52.40	50.95	52.03	11911900	-1.39	0.46
25/01/2018	51.22	51.77	51.09	51.63	16225600	-0.46	-1.39
26/01/2018	49.05	49.21	47.84	49.01	51851700	-4.23	-0.46
29/01/2018	48.23	49.36	48.14	48.68	18899900	-1.67	-4.23
30/01/2018	48.38	48.67	48.00	48.18	14341200	0.30	-1.67
31/01/2018	48.06	48.60	47.96	48.41	13118400	-0.66	0.30

Ο Πίνακας 3.7 παρουσιάζει ένα ενδεικτικό δείγμα των δεδομένων ειδήσεων για τη μετοχή της Starbucks (SBUX), στα οποία εφαρμόστηκε ανάλυση συναισθήματος με χρήση του μοντέλου Llama 3. Κάθε τίτλος ειδήσεων αξιολογείται από το μοντέλο και λαμβάνει ένα σκορ συναισθήματος ανά τίτλο (Positive = 1, Neutral = 0, Negative = -1). Στη συνέχεια, τα σκορ των μεμονωμένων τίτλων συνδυάζονται και υπολογίζεται το ημερήσιο σκορ συναισθήματος, το οποίο αντιπροσωπεύει το συνολικό επενδυτικό κλίμα για την ημέρα.

**Πίνακας 3.7:** Δεδομένα ειδήσεων για SBUX (Starbucks), ανάλυση συναισθήματος από Llama 3 και υπολογισμός ημερήσιου score συναισθήματος (ενδεικτικό δείγμα του DataFrame)

Ημερομηνία	Τίτλος	Ανάλυση Συναισθήματος από Llama 3	Σκορ ανά τίτλο	Ημερήσιο Σκορ Συναισθήματος
26/01/2018	Starbucks Falls After Q1 Report: What Analysts Are Saying	Negative	-1	(-1+0+1+1+1-1+1+1+1)/9= 0.44444
	41 Stocks Moving In Friday's Mid-Day Session	Neutral	0	
	Goldman Sachs Removes Starbucks From Conviction List, Maintains Buy Rating	Positive	1	
	Slowing Down? Economy Brakes Slightly In Q4 But Earnings Keep Impressing	Positive	1	
	Benzinga Pro's 6 Stocks To Watch Today	Positive	1	
	Goldman Removes Starbucks From Conviction Buy List As The Recent Quarter 'Contributes To A Narrative Of Inconsistency', Says Move Reflects View That There Are Risks To Back-Half Weighted Guidance Numbers	Negative	-1	
	21 Stocks Moving In Friday's Pre-Market Session	Positive	1	
	Starbucks Shares Down 5% Following Q1 Beat, Better-Than-Expected FY18 Guidance: Sees Adj. EPS \$2.48-\$2.53 vs \$2.28 Est.	Positive	1	
10 Stocks To Watch For January 26, 2018	Positive	1		

Μετά την εξαγωγή και επεξεργασία των δεδομένων, σχηματίζονται δύο σύνολα δεδομένων (DataFrames): ένα με τα χρηματιστηριακά δεδομένα και ένα με τα ημερήσια scores συναισθήματος από τις ειδήσεις. Τα δύο αυτά σετ συμπύσσονται (merge) με βάση την ημερομηνία, ώστε να δημιουργηθεί ένα ενιαίο σύνολο χαρακτηριστικών για την εκπαίδευση των μοντέλων μηχανικής μάθησης. Στην εκπαίδευση χρησιμοποιούνται το σκορ συναισθήματος της προηγούμενης μέρας για την πρόβλεψη της επόμενης, ο όγκος συναλλαγών και η ποσοστιαία μεταβολή της τιμής μετοχής της προηγούμενης ημέρας. Για τα μοντέλα που επηρεάζονται από τη διαφορά μεγέθους των χαρακτηριστικών, όπως τα Support Vector Regression (SVR) και τα Multi-Layer Perceptron (MLP), ο όγκος συναλλαγών κλιμακώνεται με διαίρεση με 1.000.000 (scaling), ώστε να διευκολύνεται η εκπαίδευση και να βελτιώνεται η σταθερότητα των αποτελεσμάτων.

Ο Πίνακας 3.8 παρουσιάζει ένα ενδεικτικό παράδειγμα της συγχώνευσης (merge) των βασικών δεδομένων που προέρχονται από τους Πίνακες 3.6 και 3.7. Συγκεκριμένα, συνδυάζονται τα χρηματιστηριακά δεδομένα της μετοχής με τα ημερήσια scores συναισθήματος από τις ειδήσεις, με βάση την ημερομηνία. Το αποτέλεσμα είναι ένα ενιαίο σύνολο χαρακτηριστικών, το οποίο περιλαμβάνει το ημερήσιο score συναισθήματος, την ποσοστιαία ημερήσια μεταβολή της τιμής, την προηγούμενη ποσοστιαία μεταβολή και τον όγκο συναλλαγών. Το σύνολο αυτό χρησιμοποιείται ως είσοδος για τα μοντέλα μηχανικής μάθησης, επιτρέποντας την εκπαίδευση και την πρόβλεψη της επόμενης ημέρας με βάση τα διαθέσιμα χαρακτηριστικά.

**Πίνακας 3.8:** Συγχώνευση (merge) των βασικών δεδομένων των Πινάκων 3.6 and 3.7 (ενδεικτικό δείγμα του merged DataFrame) για πρόβλεψη

Ημερομηνία	Ημερήσιο Σκορ Συναισθήματος	Ποσοστιαία Ημερήσια Μεταβολή Τιμής	Προηγούμενη Ποσοστιαία Ημερήσια Μεταβολή Τιμής	Όγκος
22/01/2018	0	0.245	0.278	11.9458
24/01/2018	1	-1.394	0.456	11.9119
25/01/2018	0.8	-0.460	-1.394	16.2256
26/01/2018	0.444	-4.228	-0.460	51.8517
29/01/2018	0.857	-1.673	-4.228	18.8999

Είναι σημαντικό να σημειωθεί ότι ο όγκος των δεδομένων πρέπει να είναι αρκετά μεγάλος, καλύπτοντας αρκετές ημέρες και μήνες, ώστε να επιτρέπει την αποτελεσματική εκπαίδευση των αλγορίθμων μηχανικής μάθησης. Οι πίνακες που παρουσιάζονται προηγουμένως αφορούν μόνο μερικές ημέρες και χρησιμοποιούνται αποκλειστικά ως ενδεικτικό δείγμα για την κατανόηση της διαδικασίας και της δομής των δεδομένων.

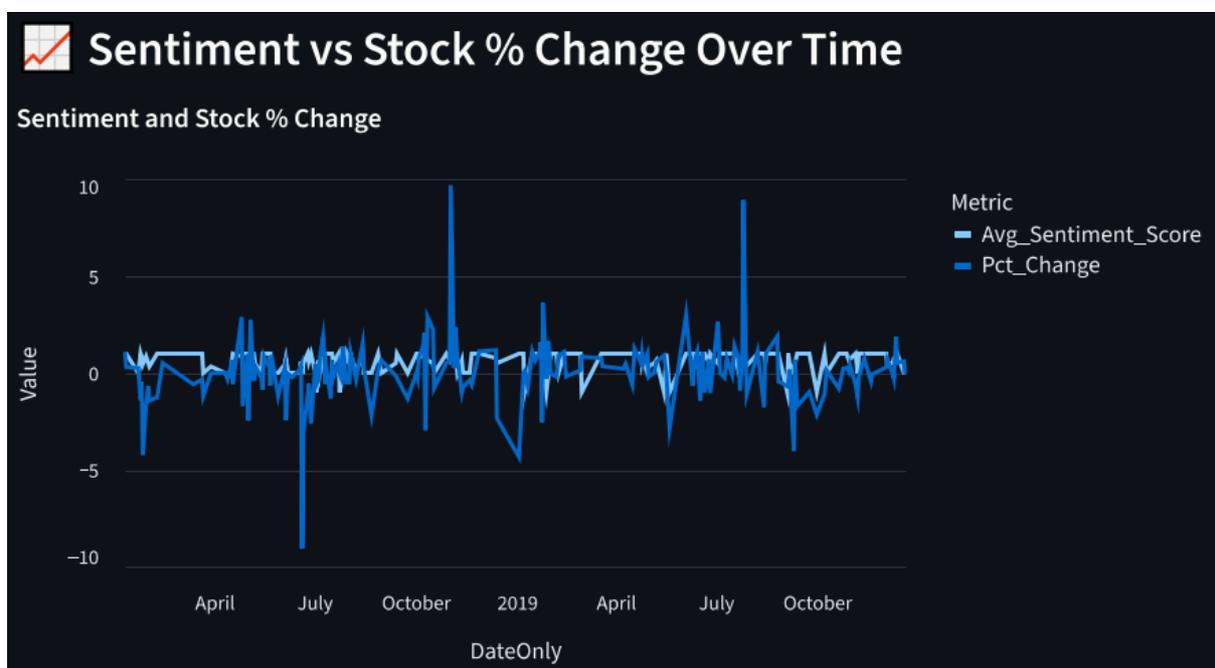
Κατά τη διαδικασία συγχώνευσης των δεδομένων, πραγματοποιήθηκε επίσης ο υπολογισμός της συσχέτισης (correlation) μεταξύ του ημερήσιου score συναισθήματος (Avg\_Sentiment\_Score) και της ποσοστιαίας μεταβολής της τιμής μετοχής (Pct\_Change). Αυτό επιτρέπει την αρχική διερεύνηση της σχέσης μεταξύ του επενδυτικού συναισθήματος και της απόδοσης της μετοχής, παρέχοντας πολύτιμη πληροφορία για την πρόβλεψη των χρηματιστηριακών μεταβολών.

Στο παρακάτω απόσπασμα κώδικα υπολογίζεται η συσχέτιση μεταξύ του ημερήσιου σκορ συναισθήματος (Avg\_Sentiment\_Score) και της ποσοστιαίας μεταβολής της τιμής μετοχής (Pct\_Change) χρησιμοποιώντας τη μέθοδο corr() της pandas. Το αποτέλεσμα εμφανίζεται μέσω του streamlit widget st.metric, ενώ στη συνέχεια δημιουργείται μια γραφική απεικόνιση της σχέσης αυτής με την εντολή st.subheader.

```
# --- Correlation metric ---
correlation = merged_data['Avg_Sentiment_Score'].corr(merged_data['Pct_Change'])
st.metric("📊 Correlation (Sentiment vs % Change)", f"{correlation:.2f}")

# --- Visualization: Sentiment vs Stock % Change ---
st.subheader("📈 Sentiment vs Stock % Change Over Time")
```

Το Σχήμα 3.6 παρουσιάζει τη γραφική απεικόνιση της σχέσης μεταξύ του ημερήσιου σκορ συναισθήματος (Avg\_Sentiment\_Score) και της ποσοστιαίας μεταβολής της τιμής μετοχής (Pct\_Change) για ένα ενδεικτικό χρονικό διάστημα. Η γραφική παράσταση επιτρέπει την οπτική αξιολόγηση της συσχέτισης μεταξύ της διάθεσης της αγοράς και των χρηματιστηριακών μεταβολών. Από το Σχήμα φαίνεται πώς οι θετικές ή αρνητικές τάσεις στο συναίσθημα μπορεί να αντιστοιχούν σε ανοδικές ή καθοδικές μεταβολές των τιμών, παρέχοντας μια αρχική ένδειξη της επίδρασης του συναισθήματος στην απόδοση των μετοχών.



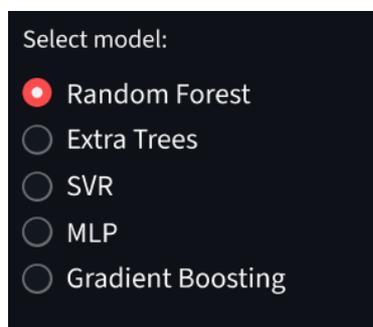
**Σχήμα 3.6:** Ημερήσιο Σκορ Συναισθήματος (Avg\_Sentiment\_Score) και Ποσοστιαία Ημερήσια Μεταβολή Τιμής (Pct\_Change)

Το αποτέλεσμα του υπολογισμού της συσχέτισης μεταξύ του ημερήσιου σκορ συναισθήματος (Avg\_Sentiment\_Score) και της ποσοστιαίας μεταβολής της τιμής μετοχής (Pct\_Change) ήταν **0,13**, τιμή που αντιστοιχεί σε μια ασθενή γραμμική σχέση (όπου 0 σημαίνει καμία γραμμική σχέση, 1 τέλεια θετική και -1 τέλεια αρνητική σχέση). Το αποτέλεσμα αυτό δείχνει ότι το συναίσθημα από μόνο του δεν επαρκεί για την αξιόπιστη πρόβλεψη της απόδοσης μιας μετοχής. Για την ανάπτυξη πιο αποτελεσματικών μοντέλων πρόβλεψης απαιτείται η ενσωμάτωση και άλλων χαρακτηριστικών, όπως ιστορικές τιμές, όγκος συναλλαγών και άλλες χρηματιστηριακές μεταβλητές, οι οποίες θα χρησιμοποιηθούν στο επόμενο στάδιο της ανάλυσης.

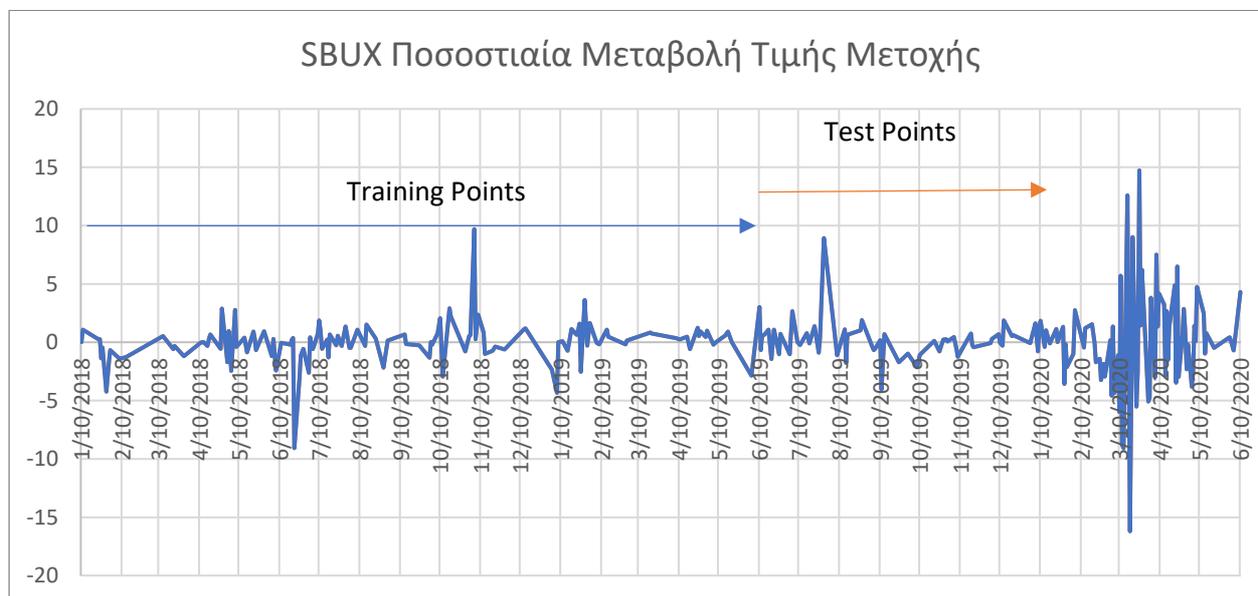
📊 Correlation (Sentiment vs % Change)  
**0.13**

### (3.4.β Εκπαίδευση Αλγορίθμων Μηχανικής Μάθησης)

Για την πρόβλεψη απόδοσης μετοχής επιλέχθηκαν οι παρακάτω αλγόριθμοι μηχανικής μάθησης που θα παρουσιαστούν αναλυτικά σε επόμενες παραγράφους.



Από τις 266 γραμμές δεδομένων που προέκυψαν από τη συγχώνευση των χρηματιστηριακών δεδομένων με τα ημερήσια σκορ συναισθήματος, το 70% χρησιμοποιήθηκε για την εκπαίδευση των αλγορίθμων, επιλέγοντας τα πρώτα χρονολογικά σημεία και όχι τυχαία, ενώ το 30% χρησιμοποιήθηκε για το τεστ, αντιστοιχώντας σε 80 σημεία. Στο Σχήμα 3.7 παρουσιάζονται τα δεδομένα εκπαίδευσης και τεστ, όπου παρατηρείται ότι τα σημεία τεστ περιλαμβάνουν θόρυβο την άνοιξη του 2020, συμβατό με την αρχή της πανδημίας COVID-19. Για τον λόγο αυτό, το συγκεκριμένο χρονικό διάστημα εξαιρέθηκε από το σετ δεδομένων. Επιπλέον, από το υπόλοιπο σετ αφαιρέθηκαν ακραίες τιμές (outliers) με τιμές μεγαλύτερες από 4 ή μικρότερες από -4. Το τελικό σετ δεδομένων περιλαμβάνει 126 σημεία για εκπαίδευση και 55 σημεία για τεστ.



Σχήμα 3.7: Ποσοστιαία Μεταβολή Τιμής Μετοχής (Pct\_Change)

Σε επόμενες παραγράφους παρατίθενται τα αποτελέσματα που προέκυψαν για κάθε χρησιμοποιημένη μέθοδο μηχανικής μάθησης. Η ανάλυση και συζήτηση αυτών των αποτελεσμάτων παρουσιάζεται στο Κεφάλαιο 4, όπου εξετάζεται η απόδοση των αλγορίθμων στις προβλέψεις της απόδοσης των μετοχών.

## Περιγραφή Κώδικα

Ο Python κώδικας της εφαρμογής αναπτύχθηκε για την εκπαίδευση και αξιολόγηση μοντέλων μηχανικής μάθησης με στόχο την πρόβλεψη της ποσοστιαίας μεταβολής της τιμής μετοχής.

Στην αρχή του προγράμματος εισάγονται όλες οι απαραίτητες βιβλιοθήκες, όπως το pandas για τη διαχείριση των δεδομένων, η scikit-learn για την υλοποίηση των μοντέλων Μηχανικής Μάθησης και την αξιολόγησή τους, καθώς και τα seaborn και matplotlib για τη δημιουργία γραφημάτων. Επιπλέον, ενεργοποιείται η βιβλιοθήκη streamlit ώστε να οριστεί το περιβάλλον της εφαρμογής.

```
import streamlit as st
import pandas as pd
from sklearn.ensemble import RandomForestRegressor, ExtraTreesRegressor, GradientBoostingRegressor
from sklearn.svm import SVR
from sklearn.neural_network import MLPRegressor
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
import seaborn as sns
import matplotlib.pyplot as plt
```

Η πρώτη ενότητα του κώδικα περιλαμβάνει βοηθητικές συναρτήσεις που είναι απαραίτητες για την προετοιμασία των δεδομένων και την εκπαίδευση των μοντέλων. Η συνάρτηση `time_series_split` έχει ως στόχο να διαχωρίσει τα δεδομένα σε σύνολο εκπαίδευσης (train) και ελέγχου (test), χρησιμοποιώντας αναλογική κατανομή 70% – 30%. Σε αντίθεση με τον τυπικό τυχαίο διαχωρισμό, εδώ διατηρείται η χρονολογική σειρά των δεδομένων.

```
def time_series_split(X, y, train_ratio=0.7):
    X = X.reset_index(drop=True)
    y = y.reset_index(drop=True)
    train_size = int(len(X) * train_ratio)
    return X.iloc[:train_size], X.iloc[train_size:], y.iloc[:train_size], y.iloc[train_size:]
```

Στη συνέχεια, η συνάρτηση `train_model` αναλαμβάνει την εκπαίδευση του μοντέλου. Δέχεται ως είσοδο μία συνάρτηση που δημιουργεί το μοντέλο, εφαρμόζει τον διαχωρισμό των δεδομένων, εκπαιδεύει το μοντέλο στο σύνολο εκπαίδευσης και παράγει προβλέψεις για το σύνολο ελέγχου, επιστρέφοντας τόσο το εκπαιδευμένο μοντέλο όσο και τις πραγματικές και προβλεπόμενες τιμές.

```
def train_model(model_class, X, y):
    X_train, X_test, y_train, y_test = time_series_split(X, y)
    model = model_class()
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)
    return model, y_test, y_pred
```

Ακολουθεί η ενότητα όπου ορίζονται τα διαθέσιμα μοντέλα μηχανικής μάθησης. Κάθε μοντέλο περιγράφεται μέσω μίας συνάρτησης με συγκεκριμένες παραμέτρους. Μεταξύ των μοντέλων περιλαμβάνονται ο Random Forest Regressor, ο Extra Trees Regressor, ο Support Vector Regressor, το νευρωνικό δίκτυο MLP και ο Gradient Boosting Regressor.

```
# Model factories
def random_forest(): return RandomForestRegressor(n_estimators=100, random_state=42)
def extra_trees(): return ExtraTreesRegressor(n_estimators=300, random_state=42)
def svr_model(): return SVR(kernel='rbf')
def mlp_model(): return MLPRegressor(hidden_layer_sizes=(50,50), max_iter=1000,
random_state=42)
def gradient_boosting(): return GradientBoostingRegressor(n_estimators=100, random_state=42)
```

Στην κύρια ενότητα της εφαρμογής ενεργοποιείται το περιβάλλον streamlit. Αρχικά εμφανίζεται τίτλος και παρέχεται η δυνατότητα στον χρήστη να ανεβάσει αρχείο CSV που περιέχει τα συγχωνευμένα δεδομένα (merged data). Μετά τη φόρτωση του αρχείου εμφανίζεται προεπισκόπηση των δεδομένων και στη συνέχεια γίνεται ο ορισμός των μεταβλητών εισόδου, οι οποίες περιλαμβάνουν τον όγκο συναλλαγών, τον ημερήσιο σκορ συναισθήματος και την προηγούμενη ποσοστιαία ημερήσια μεταβολή τιμής, καθώς και του στόχου, που είναι η ποσοστιαία μεταβολή της μετοχής.

Ο χρήστης καλείται να επιλέξει το μοντέλο εκπαίδευσης μέσα από μια λίστα επιλογών. Μόλις ενεργοποιηθεί το κουμπί της εκπαίδευσης, το πρόγραμμα δημιουργεί το αντίστοιχο μοντέλο μέσω της αντιστοίχισης που συνδέει τις επιλογές με τις συναρτήσεις μοντέλου και στη συνέχεια πραγματοποιεί την εκπαίδευση μέσω της συνάρτησης train\_model. Αμέσως μετά υπολογίζονται οι δείκτες αξιολόγησης, όπως το Mean Absolute Error, το Mean Squared Error και ο συντελεστής προσδιορισμού  $R^2$ , οι οποίοι εμφανίζονται στο περιβάλλον της εφαρμογής. Παράλληλα, δημιουργείται διάγραμμα διασποράς (scatter plot) που απεικονίζει τη σχέση μεταξύ των πραγματικών και των προβλεπόμενων τιμών, συνοδευόμενο από γραμμή τέλειας πρόβλεψης, ώστε ο χρήστης να αξιολογήσει οπτικά την ακρίβεια του μοντέλου. Τέλος, δημιουργείται ένας πίνακας που περιλαμβάνει τόσο τις πραγματικές όσο και τις προβλεπόμενες τιμές, προσφέροντας πλήρη εικόνα της απόδοσης του μοντέλου στις τελευταίες χρονικές περιόδους.

```
st.title("📈 Stock Change Prediction - Model Training")
uploaded_file = st.file_uploader("Upload merged data CSV", type=["csv"])

if uploaded_file:
    merged_data = pd.read_csv(uploaded_file)
    st.subheader("📄 Loaded Data")
    st.dataframe(merged_data.head(20))

    X = merged_data[['Volume', 'Avg_Sentiment_Score', 'Prev_Pct_Change']]
    y = merged_data['Pct_Change']

    model_choice = st.radio(
```

```

"Select model:",
["Random Forest", "Extra Trees", "SVR", "MLP", "Gradient Boosting"]
)

model_map = {
    "Random Forest": random_forest,
    "Extra Trees": extra_trees,
    "SVR": svr_model,
    "MLP": mlp_model,
    "Gradient Boosting": gradient_boosting
}

if st.button("Train Model"):
    model_class = model_map[model_choice]
    model, y_test, y_pred = train_model(model_class, X, y)

    mae = mean_absolute_error(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    r2 = r2_score(y_test, y_pred)
    st.write(f"***Mean Absolute Error (MAE):** {mae:.2f}")
    st.write(f"***Mean Squared Error (MSE):** {mse:.2f}")
    st.write(f"***R2 Score:** {r2:.2f}")

    # Scatter plot
    fig1, ax1 = plt.subplots(figsize=(8,5))
    sns.scatterplot(x=y_test, y=y_pred, color='blue', ax=ax1)
    ax1.plot([min(y_test), max(y_test)], [min(y_test), max(y_test)], 'r--', label="Perfect Prediction")
    ax1.set_xlabel("Actual % Change")
    ax1.set_ylabel("Predicted % Change")
    ax1.legend()
    st.pyplot(fig1)

    # Display predicted dataframe
    pred_df = pd.DataFrame({
        "Actual_Pct_Change": y_test,
        "Predicted_Pct_Change": y_pred
    }).reset_index(drop=True)

    st.subheader("📄 Predictions vs Actual")
    st.dataframe(pred_df)

```

## Random Forest

Η μέθοδος Random Forests αποτελεί μια ισχυρή τεχνική εκπαίδευσης συνόλου μοντέλων που χρησιμοποιείται για προβλήματα ταξινόμησης και παλινδρόμησης [8]. Η βασική ιδέα είναι η κατασκευή ενός μεγάλου αριθμού δέντρων αποφάσεων (decision trees), όπου κάθε δέντρο εκπαιδεύεται σε τυχαία υποσύνολα δεδομένων και χαρακτηριστικών. Αυτή η τυχαία επιλογή μειώνει την εξάρτηση από μεμονωμένα δεδομένα και χαρακτηριστικά, αυξάνοντας τη γενίκευση και τη σταθερότητα του μοντέλου.

Τα αποτελέσματα των μεμονωμένων δέντρων συνδυάζονται με ψηφοφορία στην περίπτωση ταξινόμησης ή μέσο όρο στην περίπτωση παλινδρόμησης, μειώνοντας την υπερπροσαρμογή (overfitting) που παρατηρείται συχνά σε μεμονωμένα δέντρα. Η διαδικασία αυτή εξασφαλίζει ότι το τελικό μοντέλο είναι πιο ανθεκτικό σε θόρυβο και σε ακραίες τιμές, ενώ διατηρεί υψηλή ακρίβεια σε σύνθετα σύνολα δεδομένων.

Ένα σημαντικό πλεονέκτημα των Random Forests είναι η δυνατότητα αξιολόγησης της σημαντικότητας των χαρακτηριστικών (feature importance), η οποία επιτρέπει την αναγνώριση των πιο καθοριστικών παραγόντων για την πρόβλεψη. Η μέθοδος έχει ευρεία εφαρμογή σε πολλούς τομείς, όπως χρηματοοικονομικές προβλέψεις, ιατρική διάγνωση, ανάλυση εικόνας, αναγνώριση προτύπων και γενικά σε προβλήματα όπου απαιτείται αξιόπιστη πρόβλεψη με πολυδιάστατα δεδομένα.

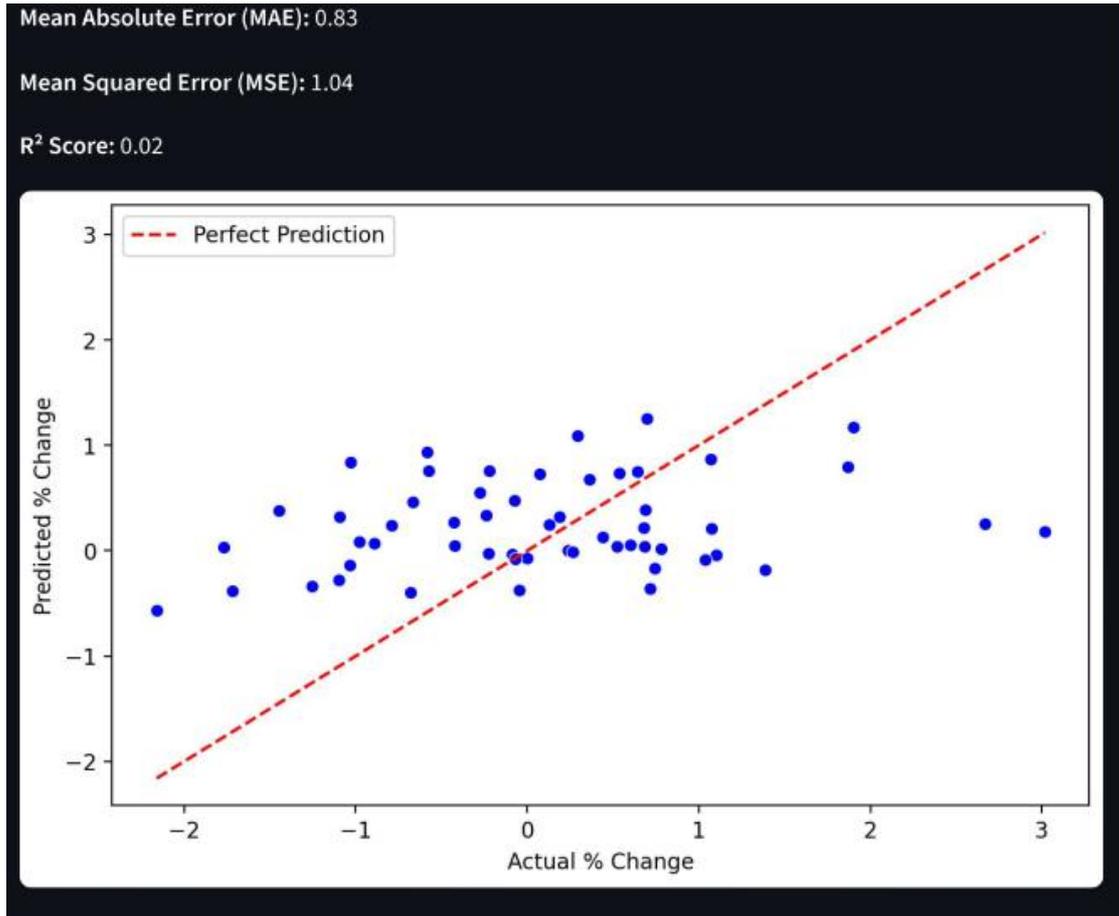
Στα Σχήματα 3.8 και 3.9 παρουσιάζονται τα αποτελέσματα της πρόβλεψης σε σημεία ελέγχου με τη χρήση του μοντέλου Random Forest.

Το Σχήμα 3.8 παρουσιάζει το διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές. Η διακεκομμένη γραμμή απεικονίζει την ιδανική κατάσταση όπου η πρόβλεψη είναι τέλεια, δηλαδή οι προβλεπόμενες τιμές συμπίπτουν με τις πραγματικές. Οι κουκκίδες δείχνουν τα σημεία πρόβλεψης του μοντέλου, δίνοντας μια οπτική εικόνα για το πόσο κοντά ή μακριά βρίσκονται από την ιδανική γραμμή.

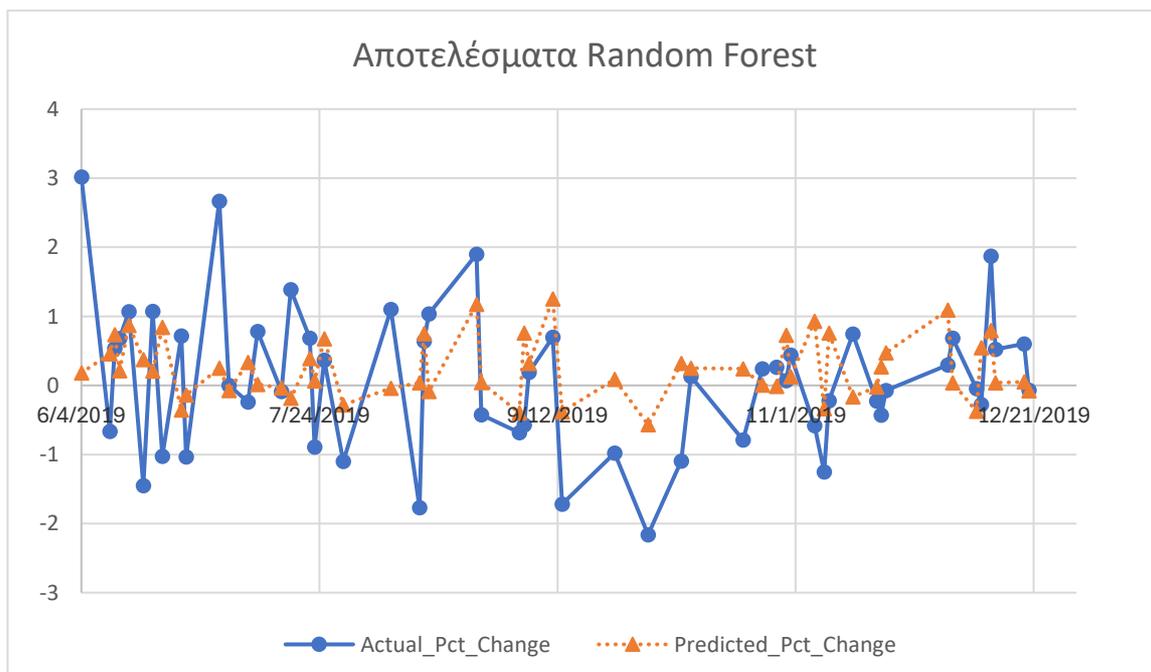
Το Σχήμα 3.9 απεικονίζει την εξέλιξη των πραγματικών (Actual\_Pct\_Change) και προβλεπόμενων (Predicted\_Pct\_Change) ποσοστιαίων αλλαγών σε σχέση με το χρόνο για το ίδιο σύνολο ελέγχου. Αυτό το γράφημα επιτρέπει την εκτίμηση της ικανότητας του μοντέλου να ακολουθεί τις χρονικές τάσεις των δεδομένων, αποκαλύπτοντας σημεία όπου η πρόβλεψη ήταν πιο ακριβής ή, αντίθετα, όπου παρουσιάστηκαν αποκλίσεις.

Η σύγκριση των δύο αυτών διαγραμμάτων δίνει μια σφαιρική εικόνα για την απόδοση του Random Forest στην πρόβλεψη των μεταβολών, παρουσιάζοντας τόσο τη συνολική ακρίβεια όσο και τη χρονική συμπεριφορά των προβλέψεων.

Για λεπτομερή συζήτηση των αποτελεσμάτων και με βάση (MAE, MSE,  $R^2$  Score) δείτε την παράγραφο 5.2.



**Σχήμα 3.8:** Διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές στα σημεία ελέγχου για Random Forest



**Σχήμα 3.9:** Σύγκριση πρόβλεψης με πραγματικές τιμές στα σημεία ελέγχου για Random Forest

## Extra Trees

Η μέθοδος Extra Trees, ή Extremely Randomized Trees, είναι μια τεχνική εκπαίδευσης συνόλου μοντέλων που χρησιμοποιείται τόσο για προβλήματα παλινδρόμησης όσο και για ταξινόμηση [10], και σχετίζεται στενά με τα Random Forests. Το κύριο χαρακτηριστικό της είναι ότι δημιουργεί μεγάλο αριθμό δέντρων αποφάσεων, αλλά αντί να επιλέγει κάθε διαχωριστικό σημείο με βάση κάποιο κριτήριο βελτιστοποίησης, επιλέγει τυχαία τα διαχωριστικά για κάθε χαρακτηριστικό, εισάγοντας μεγαλύτερη τυχαία επιλογή στο μοντέλο.

Τα Extra Trees μπορούν να εκπαιδευτούν είτε με ολόκληρο το σύνολο δεδομένων είτε με δείγματα bootstrap (τυχαία δειγματοληψία), αν και η χρήση bootstrap είναι προαιρετική. Αυτή η τυχαία διαδικασία μειώνει τη διακύμανση και την υπερπροσαρμογή (overfitting) σε σχέση με μεμονωμένα δέντρα αποφάσεων και συχνά επιτρέπει ταχύτερη εκπαίδευση σε σύγκριση με τα Random Forests, επειδή αποφεύγεται η εκτεταμένη αναζήτηση για το βέλτιστο διαχωριστικό σημείο.

Η μέθοδος είναι ιδιαίτερα αποτελεσματική σε υψηλής διάστασης δεδομένα και λειτουργεί συγκεντρώνοντας τις προβλέψεις από πολλά δέντρα για να βελτιώσει τη γενίκευση. Η κύρια διαφορά από τα Random Forests είναι η αυξημένη τυχαία επιλογή διαχωριστικών, η οποία μπορεί να αυξήσει τη σταθερότητα του μοντέλου και να μειώσει την ευαισθησία σε μικρές μεταβολές στα δεδομένα εκπαίδευσης.

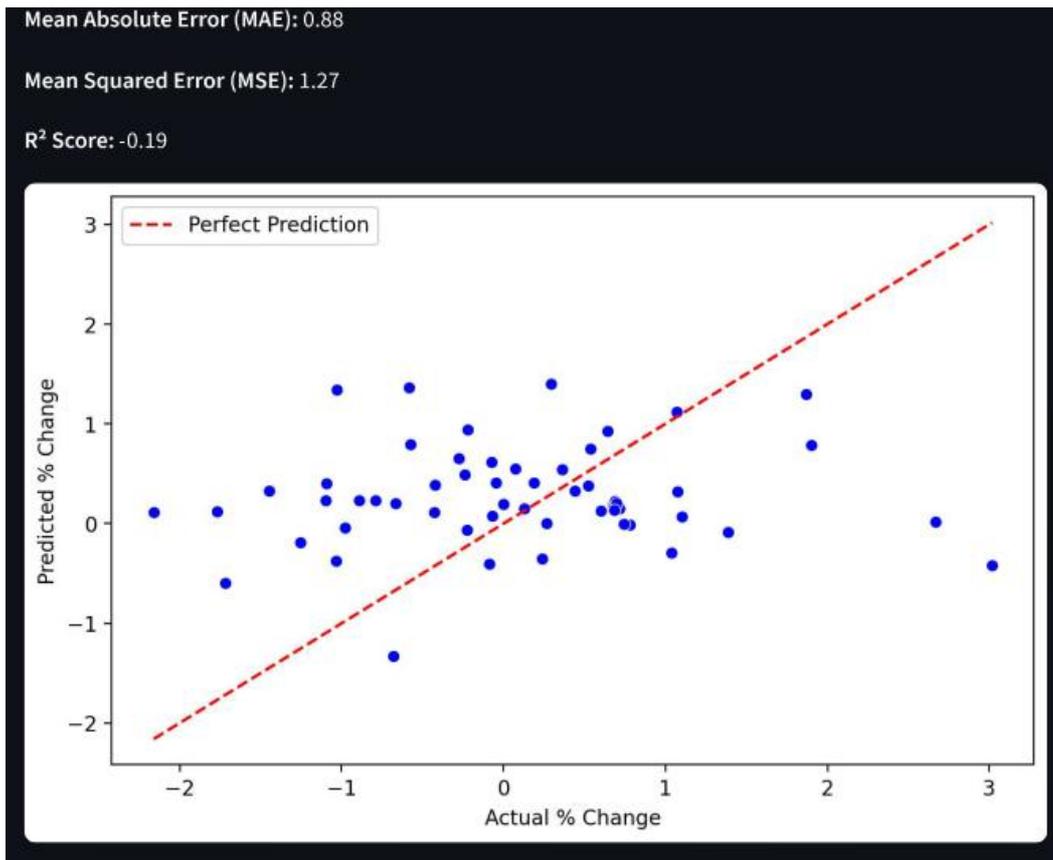
Στα Σχήματα 3.10 και 3.11 παρουσιάζονται τα αποτελέσματα της πρόβλεψης σε σημεία ελέγχου με τη χρήση του μοντέλου Extra Trees.

Το Σχήμα 3.10 παρουσιάζει το διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές. Η διακεκομμένη γραμμή απεικονίζει την ιδανική κατάσταση όπου η πρόβλεψη είναι τέλεια, δηλαδή οι προβλεπόμενες τιμές συμπίπτουν με τις πραγματικές. Οι κουκκίδες δείχνουν τα σημεία πρόβλεψης του μοντέλου, δίνοντας μια οπτική εικόνα για το πόσο κοντά ή μακριά βρίσκονται από την ιδανική γραμμή.

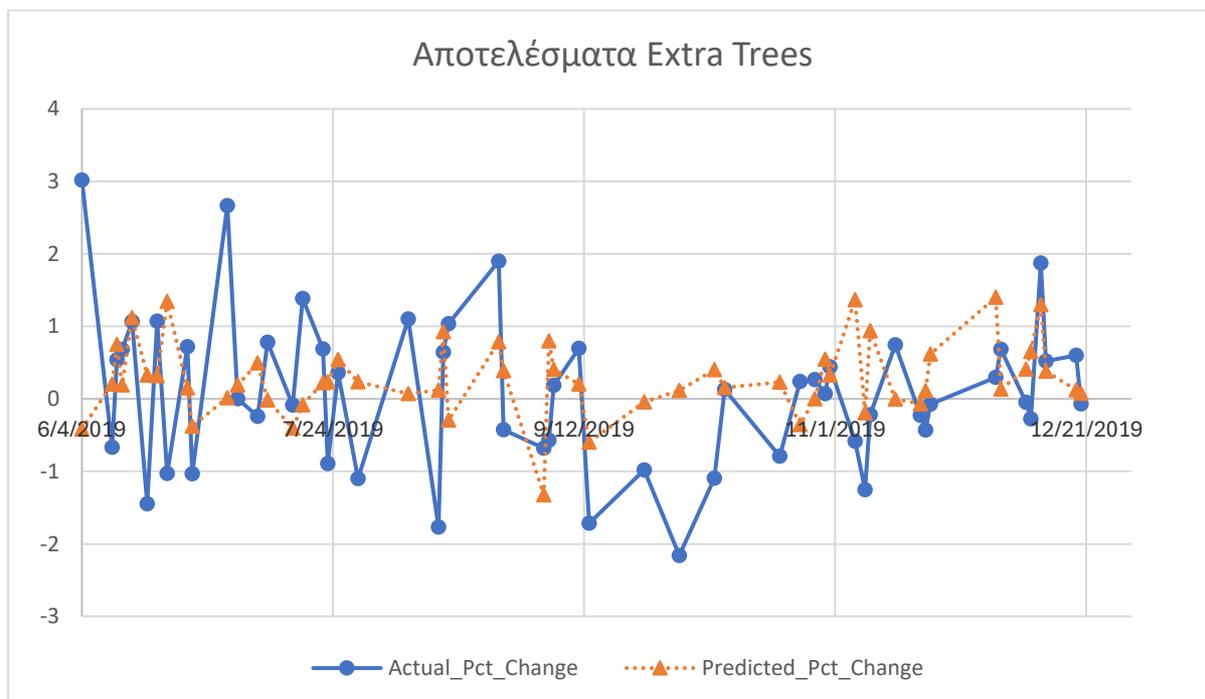
Το Σχήμα 3.11 απεικονίζει την εξέλιξη των πραγματικών (Actual\_Pct\_Change) και προβλεπόμενων (Predicted\_Pct\_Change) ποσοστιαίων αλλαγών σε σχέση με το χρόνο για το ίδιο σύνολο ελέγχου. Αυτό το γράφημα επιτρέπει την εκτίμηση της ικανότητας του μοντέλου να ακολουθεί τις χρονικές τάσεις των δεδομένων, αποκαλύπτοντας σημεία όπου η πρόβλεψη ήταν πιο ακριβής ή, αντίθετα, όπου παρουσιάστηκαν αποκλίσεις.

Η σύγκριση των δύο αυτών διαγραμμάτων δίνει μια σφαιρική εικόνα για την απόδοση του Extra Trees στην πρόβλεψη των μεταβολών, παρουσιάζοντας τόσο τη συνολική ακρίβεια όσο και τη χρονική συμπεριφορά των προβλέψεων.

Για λεπτομερή συζήτηση των αποτελεσμάτων και με βάση (MAE, MSE,  $R^2$  Score) δείτε την παράγραφο 5.2.



Σχήμα 3.10: Διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές στα σημεία ελέγχου για Extra Trees



Σχήμα 3.11: Σύγκριση πρόβλεψης με πραγματικές τιμές στα σημεία ελέγχου για Extra Trees

## Support Vector Regression (SVR)

Η μέθοδος Support Vector Regression (SVR) αποτελεί την προέκταση των Support Vector Machines (SVM) για προβλήματα παλινδρόμησης [11]. Σε αντίθεση με τις κλασικές SVM που στοχεύουν στον διαχωρισμό κατηγοριών, η SVR επιδιώκει να βρει μια συνάρτηση πρόβλεψης συνεχών τιμών, διατηρώντας τα σφάλματα προβλέψεων εντός ενός προκαθορισμένου εύρους  $\epsilon$ , γνωστού ως  $\epsilon$ -αγνοούμενη ζώνη. Μόνο τα σημεία που βρίσκονται εκτός αυτού του περιθωρίου επηρεάζουν την εκπαίδευση του μοντέλου και ονομάζονται support vectors, δηλαδή τα κρίσιμα σημεία που καθορίζουν τη μορφή της συνάρτησης πρόβλεψης.

Η SVR μπορεί να χρησιμοποιήσει πυρηνικές συναρτήσεις (kernels), όπως γραμμική, πολυωνυμική ή RBF, για να χειριστεί μη γραμμικές σχέσεις μεταξύ των εισόδων και της εξόδου, επιτρέποντας την εκμάθηση πολύπλοκων συσχετίσεων σε υψηλότερες διαστάσεις. Η τεχνική αυτή είναι ιδιαίτερα χρήσιμη σε προβλήματα όπου απαιτείται υψηλή ακρίβεια πρόβλεψης με περιορισμένη ανοχή σφάλματος, ακόμα και όταν τα δεδομένα περιέχουν θόρυβο ή αβεβαιότητα.

Η SVR βρίσκει ευρεία εφαρμογή σε χρηματοοικονομικές προβλέψεις, πρόβλεψη χρονοσειρών, ανάλυση ενεργειακών δεδομένων και επιστημονική μοντελοποίηση, όπου η δυνατότητα ελέγχου του σφάλματος πρόβλεψης και η ευελιξία χειρισμού μη γραμμικών σχέσεων είναι κρίσιμη. Παράλληλα, η επιλογή κατάλληλου πυρήνα και των υπερπαραμέτρων όπως η τιμή του  $\epsilon$ , η σταθερά C και οι παράμετροι του πυρήνα, επηρεάζουν σημαντικά την ακρίβεια και την γενίκευση του μοντέλου

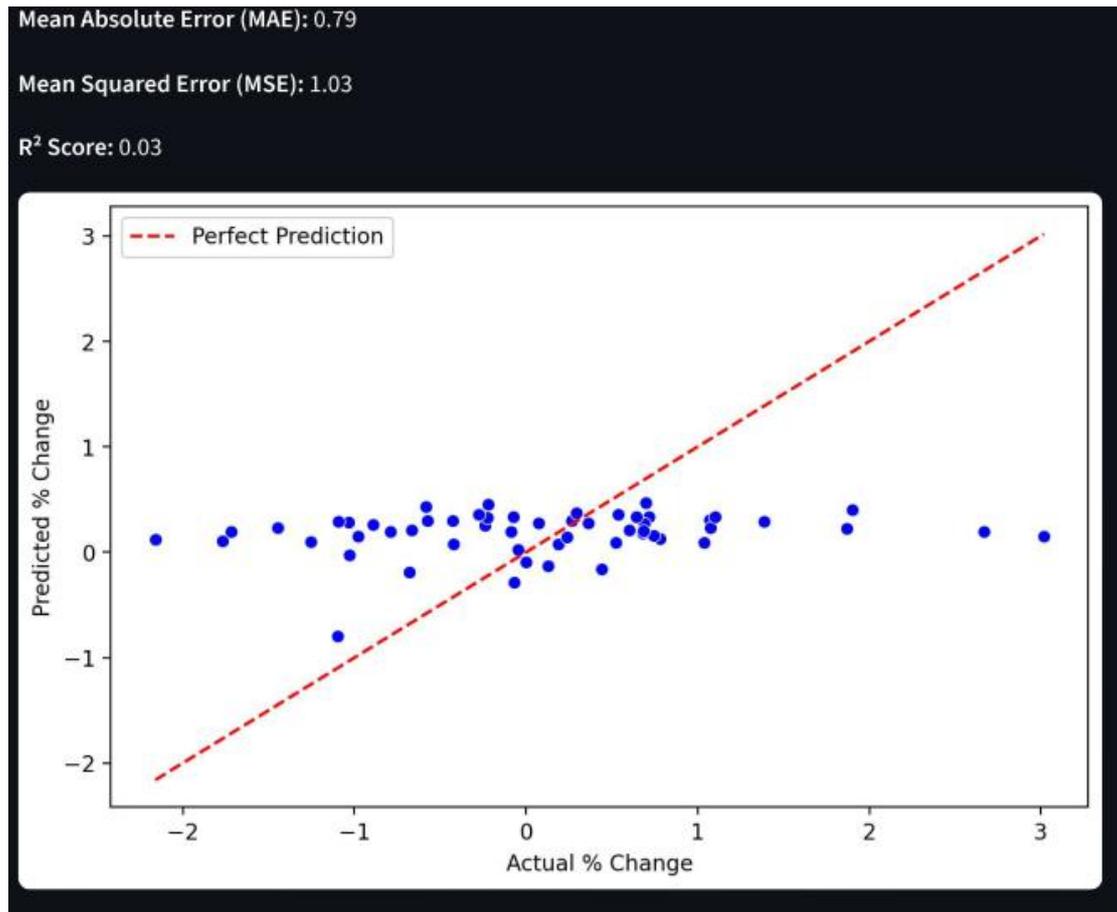
Στα Σχήματα 3.12 και 3.13 παρουσιάζονται τα αποτελέσματα της πρόβλεψης σε σημεία ελέγχου με τη χρήση του μοντέλου SVR.

Το Σχήμα 3.12 παρουσιάζει το διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές. Η διακεκομμένη γραμμή απεικονίζει την ιδανική κατάσταση όπου η πρόβλεψη είναι τέλεια, δηλαδή οι προβλεπόμενες τιμές συμπίπτουν με τις πραγματικές. Οι κουκκίδες δείχνουν τα σημεία πρόβλεψης του μοντέλου, δίνοντας μια οπτική εικόνα για το πόσο κοντά ή μακριά βρίσκονται από την ιδανική γραμμή.

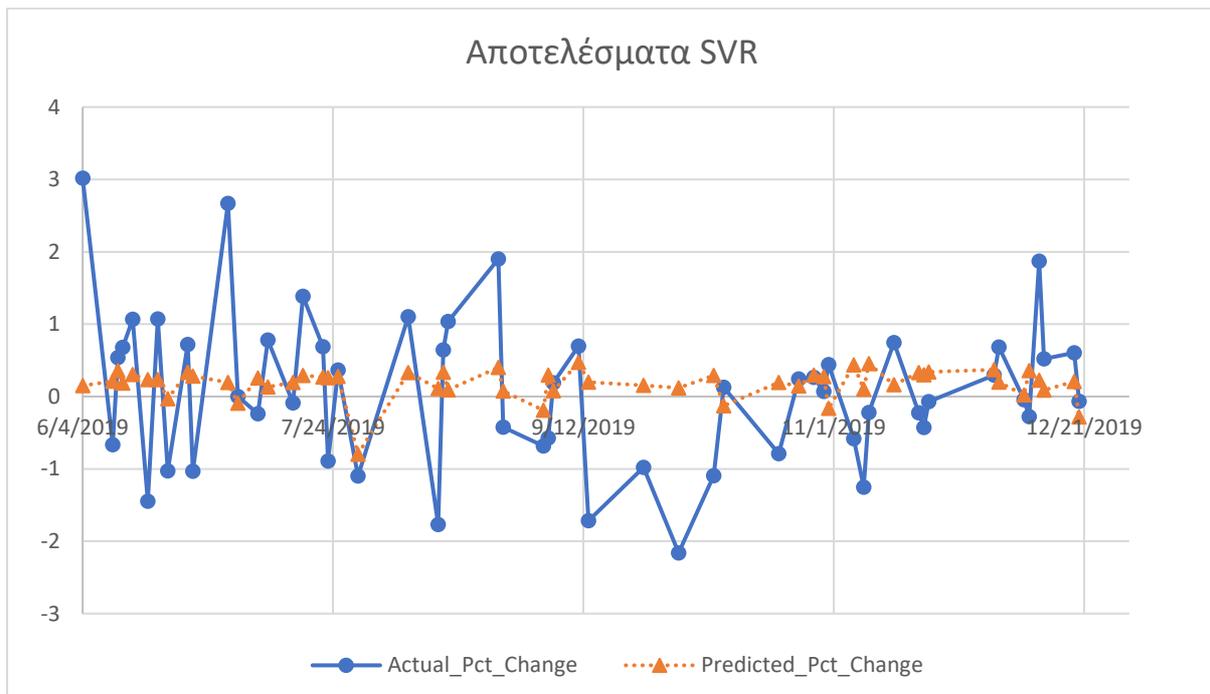
Το Σχήμα 3.13 απεικονίζει την εξέλιξη των πραγματικών (Actual\_Pct\_Change) και προβλεπόμενων (Predicted\_Pct\_Change) ποσοστιαίων αλλαγών σε σχέση με το χρόνο για το ίδιο σύνολο ελέγχου. Αυτό το γράφημα επιτρέπει την εκτίμηση της ικανότητας του μοντέλου να ακολουθεί τις χρονικές τάσεις των δεδομένων, αποκαλύπτοντας σημεία όπου η πρόβλεψη ήταν πιο ακριβής ή, αντίθετα, όπου παρουσιάστηκαν αποκλίσεις.

Η σύγκριση των δύο αυτών διαγραμμάτων δίνει μια σφαιρική εικόνα για την απόδοση του SVR στην πρόβλεψη των μεταβολών, παρουσιάζοντας τόσο τη συνολική ακρίβεια όσο και τη χρονική συμπεριφορά των προβλέψεων.

Για λεπτομερή συζήτηση των αποτελεσμάτων και με βάση (MAE, MSE,  $R^2$  Score) δείτε την παράγραφο 5.2.



Σχήμα 3.12: Διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές στα σημεία ελέγχου για SVR



Σχήμα 3.13: Σύγκριση πρόβλεψης με πραγματικές τιμές στα σημεία ελέγχου για SVR

## Multilayer Perceptron (MLP)

Η μέθοδος Multilayer Perceptron (MLP) είναι μια κλασική αρχιτεκτονική τεχνητού νευρωνικού δικτύου που χρησιμοποιείται τόσο για ταξινόμηση όσο και για παλινδρόμηση [9]. Αποτελείται από τρία βασικά είδη στρωμάτων: είσοδο, ένα ή περισσότερα κρυφά στρώματα (hidden layers) και έξοδο. Κάθε νευρώνας εφαρμόζει μια μη γραμμική συνάρτηση ενεργοποίησης (activation function), όπως sigmoid, tanh ή ReLU, η οποία επιτρέπει στο δίκτυο να μαθαίνει και να μοντελοποιεί πολύπλοκες μη γραμμικές σχέσεις.

Τα βάρη των συνδέσεων μεταξύ των νευρώνων προσαρμόζονται μέσω της μεθόδου backpropagation, η οποία χρησιμοποιεί τη μερική παράγωγο του σφάλματος πρόβλεψης ως οδηγό για ενημέρωση των βαρών, μειώνοντας σταδιακά το συνολικό σφάλμα του μοντέλου. Η διαδικασία αυτή συνδυάζεται με τεχνικές όπως η στοχαστική καθοδική βελτιστοποίηση (stochastic gradient descent) και η χρήση παραμέτρων όπως ο ρυθμός μάθησης (learning rate), οι οποίες επηρεάζουν την ταχύτητα και τη σταθερότητα της εκπαίδευσης.

Η MLP είναι ιδιαίτερα ευέλικτη και ισχυρή επειδή μπορεί να μάθει σύνθετες σχέσεις μεταξύ εισόδου και εξόδου, ακόμα και όταν αυτές είναι μη γραμμικές ή πολυδιάστατες. Έχει ευρεία εφαρμογή σε διάφορους τομείς, όπως αναγνώριση προτύπων, ταξινόμηση εικόνας, επεξεργασία φυσικής γλώσσας (NLP), πρόβλεψη χρονοσειρών, ιατρική διάγνωση και οικονομικές προβλέψεις. Μπορεί επίσης να ενσωματωθεί με τεχνικές regularization, όπως L1/L2 penalties και dropout, για να αποφευχθεί η υπερπροσαρμογή (overfitting) σε μικρά ή θορυβώδη σύνολα δεδομένων. Η αρχιτεκτονική των κρυφών στρωμάτων και η επιλογή των νευρώνων σε κάθε στρώμα επηρεάζουν σημαντικά την απόδοση και την ικανότητα γενίκευσης του δικτύου.

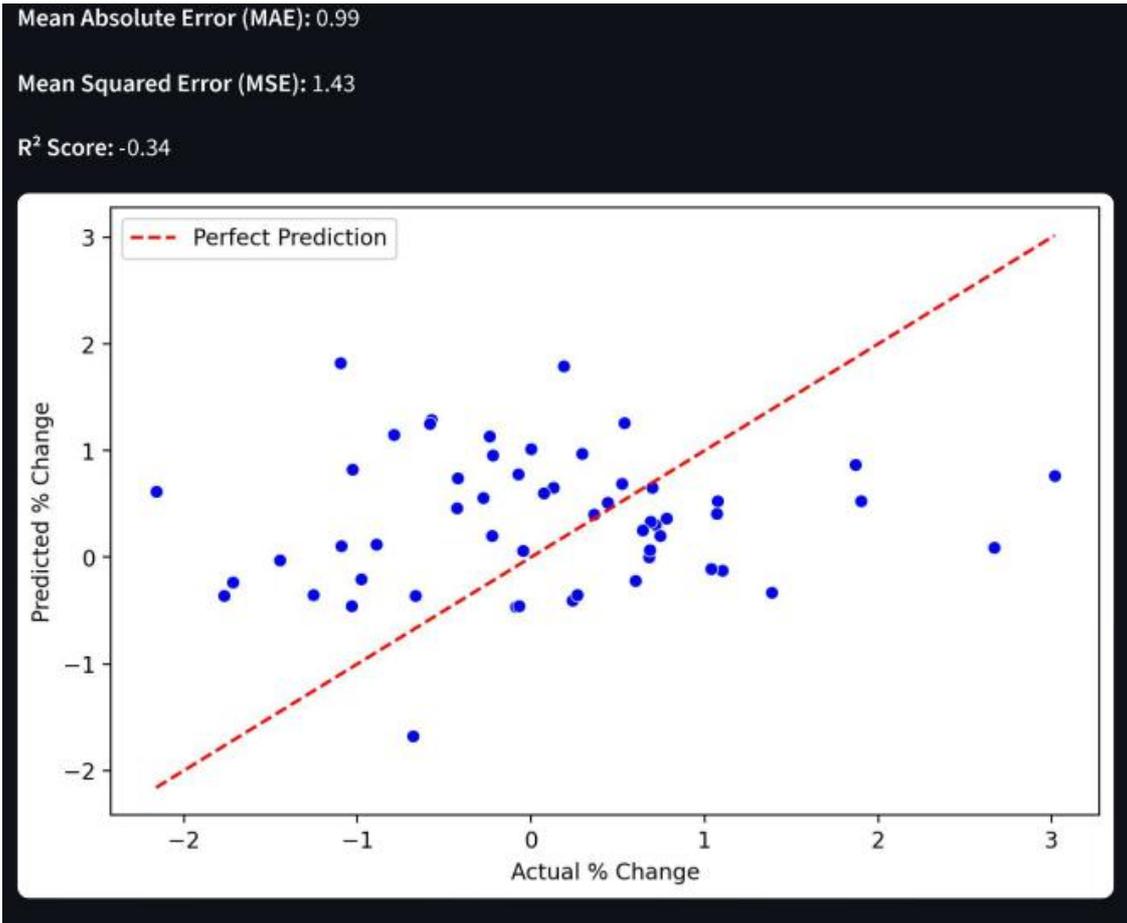
Στα Σχήματα 3.14 και 3.15 παρουσιάζονται τα αποτελέσματα της πρόβλεψης σε σημεία ελέγχου με τη χρήση του μοντέλου MLP.

Το Σχήμα 3.14 παρουσιάζει το διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές. Η διακεκομμένη γραμμή απεικονίζει την ιδανική κατάσταση όπου η πρόβλεψη είναι τέλεια, δηλαδή οι προβλεπόμενες τιμές συμπίπτουν με τις πραγματικές. Οι κουκκίδες δείχνουν τα σημεία πρόβλεψης του μοντέλου, δίνοντας μια οπτική εικόνα για το πόσο κοντά ή μακριά βρίσκονται από την ιδανική γραμμή.

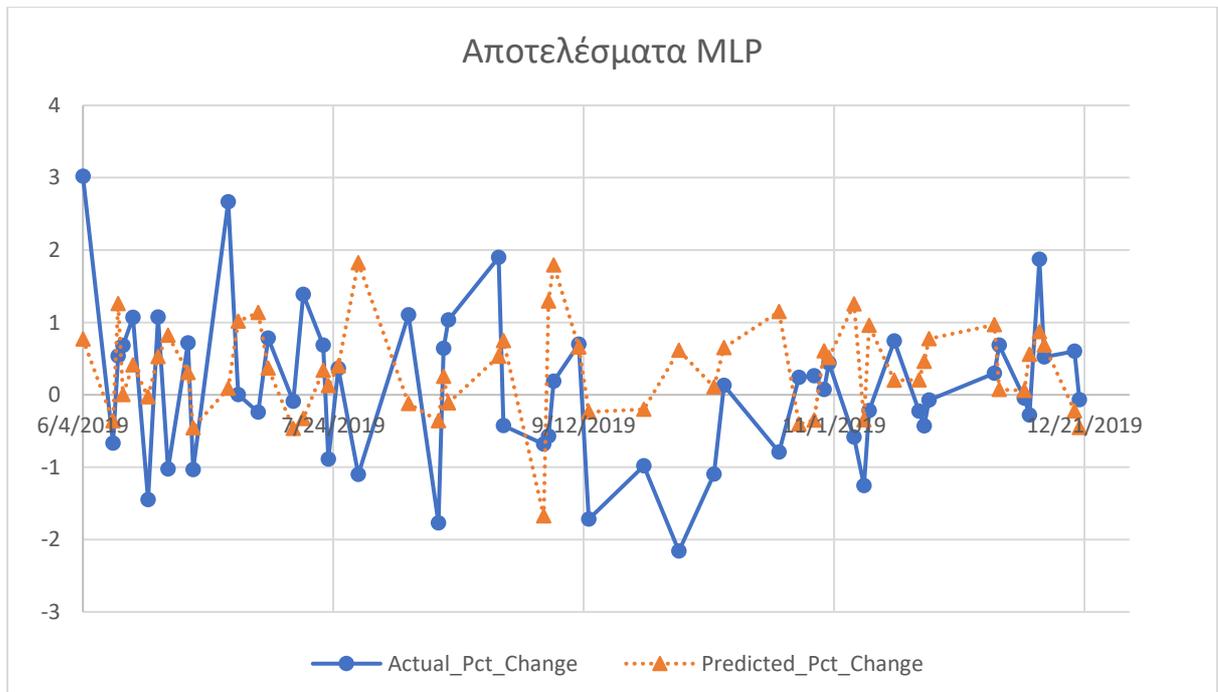
Το Σχήμα 3.15 απεικονίζει την εξέλιξη των πραγματικών (Actual\_Pct\_Change) και προβλεπόμενων (Predicted\_Pct\_Change) ποσοστιαίων αλλαγών σε σχέση με το χρόνο για το ίδιο σύνολο ελέγχου. Αυτό το γράφημα επιτρέπει την εκτίμηση της ικανότητας του μοντέλου να ακολουθεί τις χρονικές τάσεις των δεδομένων, αποκαλύπτοντας σημεία όπου η πρόβλεψη ήταν πιο ακριβής ή, αντίθετα, όπου παρουσιάστηκαν αποκλίσεις.

Η σύγκριση των δύο αυτών διαγραμμάτων δίνει μια σφαιρική εικόνα για την απόδοση του MLP στην πρόβλεψη των μεταβολών, παρουσιάζοντας τόσο τη συνολική ακρίβεια όσο και τη χρονική συμπεριφορά των προβλέψεων.

Για λεπτομερή συζήτηση των αποτελεσμάτων και με βάση (MAE, MSE, R<sup>2</sup> Score) δείτε την παράγραφο 5.2.



Σχήμα 3.14: Διάγραμμα διασποράς των προβλεπόμενων ποσοσטיαίων μεταβολών σε σχέση με τις πραγματικές στα σημεία ελέγχου για MLP



Σχήμα 3.15: Σύγκριση πρόβλεψης με πραγματικές τιμές στα σημεία ελέγχου για MLP

## Gradient Boosting

Η μέθοδος Gradient Boosting είναι μια ισχυρή τεχνική εκπαίδευσης συνόλου μοντέλων που χρησιμοποιείται τόσο για προβλήματα ταξινόμησης όσο και για προβλήματα παλινδρόμησης [12]. Η βασική της φιλοσοφία είναι η σταδιακή κατασκευή ενός ισχυρού προγνωστικού μοντέλου μέσω της συνδυαστικής εκπαίδευσης πολλών αδύναμων μοντέλων, συνήθως δέντρων αποφάσεων (decision trees). Αντί να εκπαιδεύονται όλα τα δέντρα ανεξάρτητα, κάθε νέο δέντρο επικεντρώνεται στην ελαχιστοποίηση των σφαλμάτων (residuals) του προηγούμενου συνόλου μοντέλων. Αυτό επιτυγχάνεται χρησιμοποιώντας την κατεύθυνση της αρνητικής κλίσης (gradient) της συνάρτησης απώλειας ως οδηγό για τη διόρθωση των σφαλμάτων.

Η διαδικασία εκπαίδευσης είναι επαναληπτική: σε κάθε βήμα προστίθεται ένα νέο δέντρο που μαθαίνει να προβλέπει τα σφάλματα του προηγούμενου συνόλου, ενώ οι προβλέψεις του προηγούμενου μοντέλου ενημερώνονται κατάλληλα. Αυτή η προσέγγιση μειώνει την διακύμανση και την προκατάληψη του τελικού μοντέλου και επιτρέπει την αντιμετώπιση πολύπλοκων, μη γραμμικών σχέσεων στα δεδομένα. Το Gradient Boosting μπορεί να προσαρμοστεί εύκολα με παραμέτρους όπως ο αριθμός των δέντρων, το βάθος κάθε δέντρου, ο ρυθμός εκμάθησης (learning rate) και η συνάρτηση απώλειας, προσφέροντας μεγάλη ευελιξία ανάλογα με το πρόβλημα.

Η μέθοδος έχει βρει ευρεία εφαρμογή σε χρηματοοικονομικές προβλέψεις, ανάλυση εικόνας, πρόβλεψη χρονοσειρών, ιατρική διάγνωση και άλλα προβλήματα μηχανικής μάθησης όπου απαιτείται υψηλή ακρίβεια και δυνατότητα χειρισμού πολύπλοκων σχέσεων μεταξύ χαρακτηριστικών και στόχου.

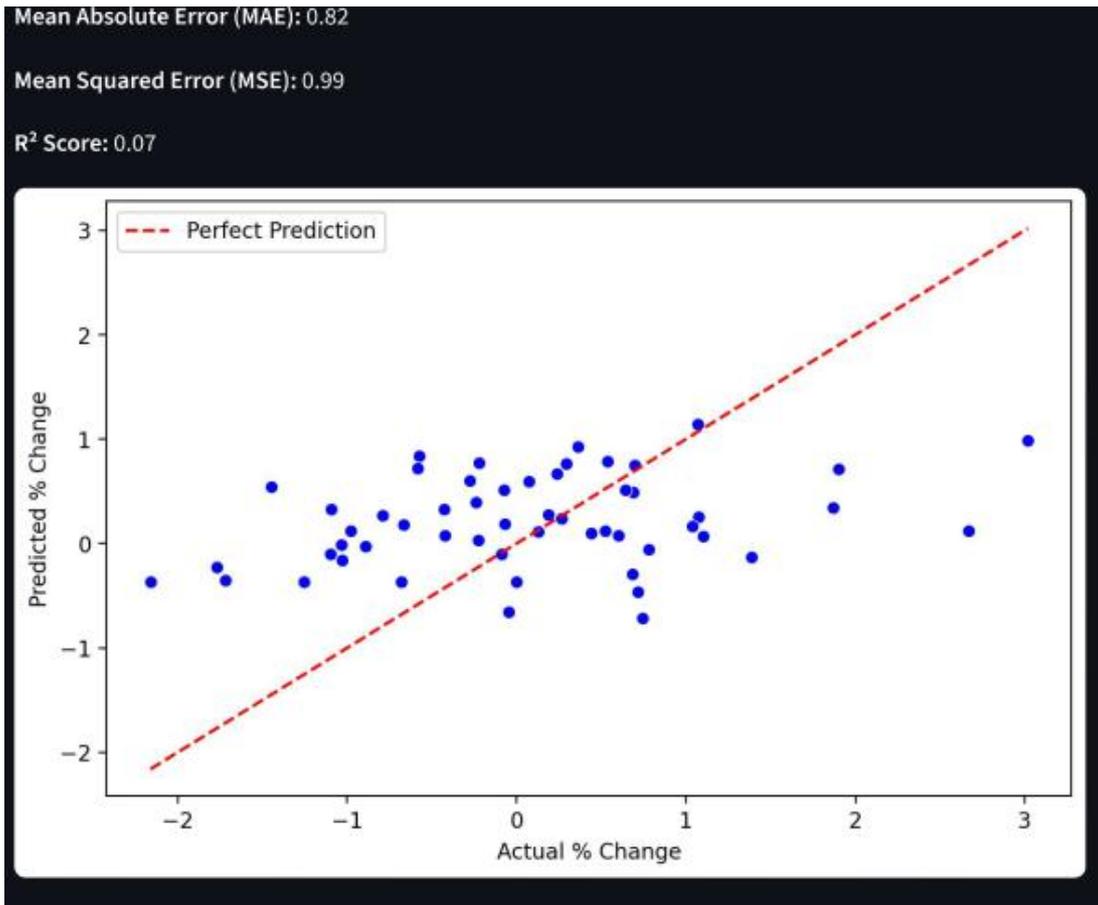
Στα Σχήματα 3.16 και 3.17 παρουσιάζονται τα αποτελέσματα της πρόβλεψης σε σημεία ελέγχου με τη χρήση του μοντέλου GB.

Το Σχήμα 3.16 παρουσιάζει το διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές. Η διακεκομμένη γραμμή απεικονίζει την ιδανική κατάσταση όπου η πρόβλεψη είναι τέλεια, δηλαδή οι προβλεπόμενες τιμές συμπίπτουν με τις πραγματικές. Οι κουκκίδες δείχνουν τα σημεία πρόβλεψης του μοντέλου, δίνοντας μια οπτική εικόνα για το πόσο κοντά ή μακριά βρίσκονται από την ιδανική γραμμή.

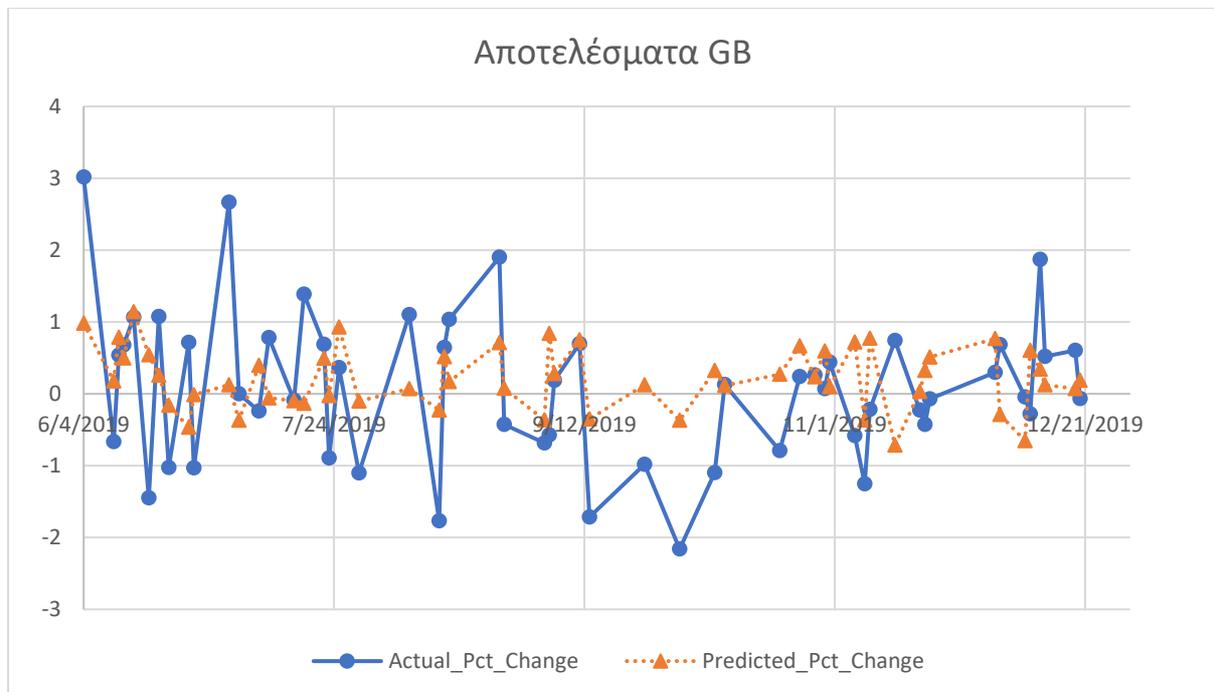
Το Σχήμα 3.17 απεικονίζει την εξέλιξη των πραγματικών (Actual\_Pct\_Change) και προβλεπόμενων (Predicted\_Pct\_Change) ποσοστιαίων αλλαγών σε σχέση με το χρόνο για το ίδιο σύνολο ελέγχου. Αυτό το γράφημα επιτρέπει την εκτίμηση της ικανότητας του μοντέλου να ακολουθεί τις χρονικές τάσεις των δεδομένων, αποκαλύπτοντας σημεία όπου η πρόβλεψη ήταν πιο ακριβής ή, αντίθετα, όπου παρουσιάστηκαν αποκλίσεις.

Η σύγκριση των δύο αυτών διαγραμμάτων δίνει μια σφαιρική εικόνα για την απόδοση του GB στην πρόβλεψη των μεταβολών, παρουσιάζοντας τόσο τη συνολική ακρίβεια όσο και τη χρονική συμπεριφορά των προβλέψεων.

Για λεπτομερή συζήτηση των αποτελεσμάτων και με βάση (MAE, MSE, R<sup>2</sup> Score) δείτε την παράγραφο 5.2.



Σχήμα 3.16: Διάγραμμα διασποράς των προβλεπόμενων ποσοστιαίων μεταβολών σε σχέση με τις πραγματικές στα σημεία ελέγχου για GB



Σχήμα 3.17: Σύγκριση πρόβλεψης με πραγματικές τιμές στα σημεία ελέγχου για GB

## (3.5 Σύνοψη & Εξαγωγή Δεδομένων Οικονομικών Αναφορών)

---

Η αυτόματη σύνοψη κειμένου (Automatic Text Summarization, ATS) εξελίχθηκε σημαντικά με την πάροδο του χρόνου [14]. Στα πρώτα στάδια, πριν το 2000, οι μέθοδοι βασίζονταν κυρίως σε στατιστικά χαρακτηριστικά, όπως η συχνότητα λέξεων και η θέση των προτάσεων, με περιορισμένη κατανόηση συμφραζομένων. Μεταξύ 2000 και 2010, οι μέθοδοι μηχανικής μάθησης και τα επιφανειακά νευρωνικά δίκτυα επέτρεψαν την αυτοματοποιημένη εξαγωγή χαρακτηριστικών, ενώ η απόδοσή τους εξαρτιόταν από χειροποίητα ειδικά στοιχεία. Στην περίοδο 2010 – 2020, η ανάπτυξη μεθόδων βαθιάς μάθησης βελτίωσε την ποιότητα των εξαγωγικών και αφαιρετικών συνοψίσεων, επιτρέποντας καλύτερη μοντελοποίηση των συμφραζομένων, παρότι εμφανίστηκαν προκλήσεις ως προς τη συνεκτικότητα και την ακρίβεια.

Η εποχή των μεγάλων γλωσσικών μοντέλων (LLM), όπως το LLaMa και το ChatGPT, έφερε σημαντική ανατροπή στην αυτόματη σύνοψη κειμένου, προσφέροντας ευελιξία στην παραγωγή συνόψεων μέσω εξαγωγικών (extractive), αφαιρετικών (abstractive) και υβριδικών (hybrid) τεχνικών. Τα LLM αξιοποιούν τη μάθηση σε πλαίσιο (in-context learning) και τη μάθηση με λίγα παραδείγματα (few-shot learning), δημιουργώντας πλούσιες και προσαρμοστικές περιλήψεις. Παρά τα πλεονεκτήματα, η χρήση τους απαιτεί προσοχή, καθώς μπορεί να προκύψουν ανακρίβειες ή δυσκολία στην τήρηση συγκεκριμένων οδηγιών, ιδιαίτερα σε εξειδικευμένα οικονομικά κείμενα.

Στην εργασία το Llama 3 χρησιμοποιήθηκε για την αυτόματη σύνοψη και ανάλυση οικονομικών αναφορών, επιτρέποντας την εξαγωγή δομημένων δεδομένων όπως οικονομικοί δείκτες, χρονικές περίοδοι και μεταβολές στα αποτελέσματα χρήσης. Η δυνατότητα του μοντέλου να επεξεργάζεται φυσική γλώσσα επιτρέπει την αποτελεσματική μείωση της πολυπλοκότητας των αναφορών και τη μετατροπή τους σε χρήσιμες πληροφορίες για περαιτέρω ανάλυση, συμβάλλοντας στη λήψη αποφάσεων και στη δημιουργία στρατηγικών βασισμένων σε δεδομένα.

### (3.5.α Τεχνική Προσέγγιση)

---

Τα προ-εκπαιδευμένα μοντέλα LLaMA αποτελούν μεγάλα γενικά γλωσσικά μοντέλα, εκπαιδευμένα σε τεράστια κείμενα, με εφαρμογές που περιλαμβάνουν περίληψη, ερωτοαπαντήσεις, μετάφραση και άλλες εργασίες επεξεργασίας φυσικής γλώσσας. Η λειτουργία τους βασίζεται στη μάθηση μέσα στο πλαίσιο (in-context learning) και στη μηχανική προτροπών (prompt engineering), επιτρέποντας την προσαρμογή τους σε συγκεκριμένες εργασίες χωρίς πλήρη επανεκπαίδευση.

Για την παραγωγή πιο αξιόπιστων και προσαρμοσμένων περιλήψεων οικονομικών αναφορών, χρησιμοποιούνται οι ακόλουθες τεχνικές:

1. Διαχωρισμός Κειμένου (Chunking): Το κείμενο χωρίζεται σε μικρότερα τμήματα (π.χ. 3000–5000 χαρακτήρων), τα οποία συνοψίζονται ξεχωριστά. Οι μεμονωμένες περιλήψεις στη συνέχεια συνδυάζονται για τη δημιουργία μιας ολοκληρωμένης περίληψης, διασφαλίζοντας την πληρότητα και τη συνεκτικότητα του αποτελέσματος.
2. Σχεδίαση Προτροπών (Prompt Engineering): Χρησιμοποιούνται ειδικά σχεδιασμένα prompts για την εξαγωγή συγκεκριμένων πληροφοριών ή τη διαμόρφωση της εξόδου.
3. Γενιά ενισχυμένη με ανάκτηση (Retrieval-Augmented Generation, RAG): Συνδυάζει το LLaMA 3 με εξωτερικές πηγές, όπως προηγούμενες οικονομικές αναφορές ή βάσεις δεδομένων, για

την άντληση σχετικών πληροφοριών. Η τεχνική αυτή βελτιώνει την ακρίβεια των περιλήψεων και μειώνει τον κίνδυνο παραλείψεων σημαντικών στοιχείων.

4. Λεπτομερής προσαρμογή (Fine-Tuning) / Προσαρμογή χαμηλής τάξης (Low-Rank Adaptation, LoRA): Το μοντέλο μπορεί να προσαρμοστεί σε δεδομένα συγκεκριμένου τομέα, όπως οικονομικές αναφορές, μέσω τεχνικών όπως η Low-Rank Adaptation (LoRA). Η προσέγγιση αυτή επιτρέπει ελαφριά προσαρμογή χωρίς την ανάγκη εκ νέου εκπαίδευσης ολόκληρου του μοντέλου, διατηρώντας παράλληλα την ισχύ και τη γενικότητα του προ-εκπαιδευμένου LLM.
5. Κανονικοποίηση εξόδου (Output Normalization): Περιλαμβάνει τον έλεγχο της μορφοποίησης αριθμών, νομισμάτων, ποσοστών και ημερομηνιών, καθώς και τη διασφάλιση λογικής συνέπειας (π.χ. καθαρά κέρδη  $\leq$  έσοδα), προκειμένου να παραχθούν ακριβή και αξιοποιήσιμα δεδομένα.

Η συνδυαστική χρήση αυτών των τεχνικών επιτρέπει στο LLaMA 3 να παράγει ολοκληρωμένες και ακριβείς περιλήψεις οικονομικών αναφορών, να εξαγάγει δομημένα δεδομένα και να υποστηρίζει την ανάλυση και τη λήψη αποφάσεων σε σύνθετα χρηματοοικονομικά περιβάλλοντα.

### (3.5.6 Περιγραφή Εφαρμογής)

---

Η εφαρμογή που αναπτύχθηκε αποσκοπεί στη δημιουργία ενός αυτοματοποιημένου συστήματος ικανού να παρέχει συνοπτική παρουσίαση των σημαντικότερων οικονομικών πληροφοριών, ενώ ταυτόχρονα εξαγάγει δομημένα δεδομένα σε μορφές κατάλληλες για περαιτέρω ποσοτική και ποιοτική ανάλυση, όπως πίνακες ή αρχεία JSON. Τα δεδομένα που εξαγονται περιλαμβάνουν βασικούς χρηματοοικονομικούς δείκτες (έσοδα, EBITDA, καθαρά κέρδη, περιθώρια, ROE), συγκρίσεις περιόδων (π.χ. Q1 2025 έναντι Q1 2024) και λεκτικές πληροφορίες, όπως δηλώσεις της διοίκησης και αναφορές παραγόντων κινδύνου.

Για την υλοποίηση της εφαρμογής αξιοποιούνται σύγχρονα εργαλεία και βιβλιοθήκες: τα LangChain και LlamaIndex για την ανάπτυξη διαδικασιών RAG (Retrieval-Augmented Generation), το PyMuPDF για την εξαγωγή κειμένου από αρχεία PDF, καθώς και οι pandas και NumPy για την επεξεργασία των δεδομένων. Ο συνδυασμός αυτών με το Llama 3 καθιστά δυνατή τη δημιουργία ενός ολοκληρωμένου συστήματος ανάλυσης οικονομικών αναφορών, το οποίο ενισχύει τόσο την ταχύτητα όσο και την αξιοπιστία της διαδικασίας λήψης αποφάσεων.

Η διαδικασία ξεκινά με την εισαγωγή του PDF μέσω streamlit και την εξαγωγή του κειμένου με χρήση PyMuPDF. Στη συνέχεια, το κείμενο διαχωρίζεται σε τμήματα (chunks) για να ανταποκριθεί στα όρια εισόδου του Llama 3, και ακολουθεί η παράλληλη σύνοψη κάθε τμήματος. Οι μεμονωμένες περιλήψεις συνενώνονται σε μία τελική συνοπτική αναφορά, από την οποία εξαγονται τα βασικά οικονομικά μεγέθη και αναλύονται στρατηγικές. Η προσέγγιση αυτή καθιστά εφικτή την αυτοματοποιημένη και αξιόπιστη ανάλυση μεγάλου όγκου οικονομικών εγγράφων, μειώνοντας σημαντικά τον χρόνο επεξεργασίας και περιορίζοντας τα πιθανά ανθρώπινα σφάλματα.

Η εφαρμογή αναπτύχθηκε με χρήση streamlit, προσφέροντας διαδραστικό περιβάλλον για σύνοψη και εξαγωγή δεδομένων από οικονομικές αναφορές σε PDF.

Αρχικά, ορίζεται ο τίτλος της σελίδας και η διάταξη, ενώ εμφανίζεται η επικεφαλίδα και η βασική περιγραφή.

```
st.set_page_config(page_title="Σύνοψη Οικονομικών Αρχείων", layout="wide")
st.title("Σύνοψη/Ανάλυση Οικονομικών Δεδομένων")
st.write("Ταχύτερη επεξεργασία μέσω βελτιστοποιημένων κλήσεων Ollama.")
```

Η εφαρμογή επιτρέπει επιλογή γλώσσας και λειτουργίας («Αναλυτική» ή «Γρήγορη»).

Γρήγορη λειτουργία: Τα κομμάτια είναι μικρότερα (5000 χαρακτήρες), επομένως η επεξεργασία είναι πιο γρήγορη, αλλά η λεπτομέρεια λιγότερη. Χρησιμοποιούνται μόνο 2 threads (max workers), περιορίζοντας την ταυτόχρονη χρήση πόρων για γρήγορη εκτίμηση.

Αναλυτική λειτουργία: Το κείμενο του PDF χωρίζεται σε μεγάλα κομμάτια των 10000 χαρακτήρων. Χρησιμοποιούνται 6 threads (maximum workers) παράλληλα για την επεξεργασία των chunks, ώστε να επιταχυνθεί η σύνοψη.

```
if mode == "Αναλυτική":
    CHUNK_SIZE = 10000
    MAX_WORKERS = 6
elif mode == "Γρήγορη":
    CHUNK_SIZE = 5000
    MAX_WORKERS = 2
```

Το PDF ανεβαίνει μέσω streamlit και το κείμενο εξάγεται με PyMuPDF. Η βιβλιοθήκη PyMuPDF ανοίγει το PDF στη μνήμη, διαβάζει κάθε σελίδα και εξάγει το περιεχόμενο σε μορφή κειμένου. Το κείμενο από όλες τις σελίδες συνενώνεται σε ένα ενιαίο string, έτοιμο για περαιτέρω επεξεργασία, όπως διαχωρισμός σε τμήματα (chunks) και δημιουργία περιλήψεων με το LLaMA 3. Η χρήση caching μέσω streamlit εξασφαλίζει ότι η ίδια διαδικασία δεν επαναλαμβάνεται σε επόμενα φορτώματα, μειώνοντας τον χρόνο επεξεργασίας.

```
@st.cache_data(show_spinner="📄 Εξαγωγή κειμένου από PDF...")
def extract_text_from_pdf(pdf_bytes):
    doc = fitz.open(stream=pdf_bytes, filetype="pdf")
    return " ".join(page.get_text() for page in doc)
```

Το κείμενο χωρίζεται σε τμήματα (chunks) για να ικανοποιήσει τα όρια εισόδου του Llama 3. Το LLaMA 3 έχει όριο εισόδου (context window) περίπου 4096 tokens, που αντιστοιχεί σε περίπου 3000–4000 λέξεις ή 12000–16000 χαρακτήρες ανάλογα με τη γλώσσα και τη σύνθεση του κειμένου. Για μεγαλύτερα κείμενα, όπως πολυσέλιδες οικονομικές αναφορές, απαιτείται chunking, δηλαδή διαχωρισμός του κειμένου σε μικρότερα τμήματα, ώστε κάθε κομμάτι να παραμένει εντός των ορίων και να επεξεργάζεται σωστά από το μοντέλο. Αυτό εξασφαλίζει ότι δεν χάνεται σημαντική πληροφορία και ότι οι συνολικές περιλήψεις είναι αξιόπιστες.

```
def chunk_text(text, chunk_size=5000):
    chunks = [text[i:i + chunk_size] for i in range(0, len(text), chunk_size)]
    return chunks
```

Κάθε chunk συνοψίζεται ξεχωριστά μέσω του Llama 3 δημιουργώντας υποπεριλήψεις για συνδυασμό.

```
def summarize_chunk(chunk, language):
    prompt = f"Δώσε μια συνοπτική περίληψη στα {language}:\n\n{chunk}"
    return Llama 3_ollama(prompt)
```

Οι υποπεριλήψεις συνδυάζονται σε μια τελική ζητούμενη περίληψη.

```
def summarize_text(text):
    start_total = time.time() # 🕒 Έναρξη συνολικής μέτρησης

    chunks = chunk_text(text, CHUNK_SIZE)

    if mode == "Γρήγορη" and len(chunks) > 3:
        chunks = chunks[:3]
        st.warning(" ⚡ Fast Mode: χρησιμοποιούνται μόνο τα πρώτα 3 chunks για γρήγορη εκτίμηση.")

    st.info(f" 🧠 Διαχωρίστηκαν σε {len(chunks)} τμήματα ({mode} mode).")

    summaries = []
    progress_bar = st.progress(0)

    with concurrent.futures.ThreadPoolExecutor(max_workers=MAX_WORKERS) as executor:
        futures = {executor.submit(summarize_chunk, chunk, language): idx for idx, chunk in enumerate(chunks)}
        for i, future in enumerate(concurrent.futures.as_completed(futures)):
            result = future.result()
            summaries.append(result)
            progress_bar.progress((i + 1) / len(chunks))
            st.write(f" ✅ Ολοκληρώθηκε τμήμα {i + 1}/{len(chunks)}")

    # Τελική συνδυαστική περίληψη
    final_prompt = (
        f"Συνδύασε τις παρακάτω περιλήψεις και γράψε μια συνολική περίληψη στα {language}:\n\n"
```

```

+ "\n\n".join(summaries)
)
final_summary = llama3_ollama(final_prompt)

total_elapsed = time.time() - start_total # 🕒 Συνολικός χρόνος
minutes = int(total_elapsed // 60)
seconds = int(total_elapsed % 60)

st.subheader("🧠 Τελική Περίληψη")
st.write(final_summary)
st.success(f"🏁 Χρόνος περίληψης: {minutes}m {seconds}s ({total_elapsed:.2f} sec)")

return final_summary

```

Επιπλέον, εξάγονται βασικά οικονομικά μεγέθη και στρατηγικές πληροφορίες μέσω ειδικών prompts και υπολογίζεται ο υπολογιστικός χρόνος.

```

if pdf_file is not None:
    # 🕒 Έναρξη συνολικού χρόνου όλων των βημάτων
    start_total = time.time()

    # Εξαγωγή κειμένου
    text = extract_text_from_pdf(pdf_file.read())
    st.write(f"Μήκος κειμένου: {len(text)} χαρακτήρες")
    st.success(f"✅ Εξήχθη κείμενο ({len(text)} χαρακτήρες)")

    # Περίληψη
    st.header("📄 Περίληψη Αναφοράς")
    summary = summarize_text(text)

    # Βασικά Οικονομικά Μεγέθη
    st.header("📊 Βασικά Οικονομικά Μεγέθη")
    metrics_prompt = f"""
Από το κείμενο, δώσε συνοπτικά:
- Πωλήσεις 9Μ24 και σύγκριση με το 2023
- EBITDA 9Μ24 και σύγκριση με το 2023
- Καθαρά Κέρδη 9Μ24 και σύγκριση με το 2023
- Απασχολούμενα κεφάλαια 9Μ24 και σύγκριση με το 2023
- Καθαρός Δανεισμός 9Μ24 και σύγκριση με το 2023
- Προσωρινό μέρισμα ανά μετοχή στους μετόχους
- Επιπτώσεις των τιμών του πετρελαίου

Γράψε στα {language}.

```

```

Κείμενο:
{text[:8000]}
"""

metrics = llama3_ollama(metrics_prompt)
st.write(metrics)

# Insights
st.header(" ⚡ Στρατηγική")
insights_prompt = f"""
Από το κείμενο, δώσε:
- Κύριες εξελίξεις γενικά και σε κάθε στρατηγικό κλάδο

Γράψε στα {language}.
Κείμενο:
{text[:8000]}
"""

insights = llama3_ollama(insights_prompt)
st.write(insights)

# 🕒 Τέλος συνολικού χρόνου
end_total = time.time()
total_elapsed = end_total - start_total

# Εμφάνιση συνολικού χρόνου (λεπτά:δευτερόλεπτα)
minutes = int(total_elapsed // 60)
seconds = int(total_elapsed % 60)
st.success(f" 🏁 Συνολικός χρόνος επεξεργασίας: {minutes}m {seconds}s ({total_elapsed:.2f}
sec)")

```

### (3.5.γ Παράδειγμα Εφαρμογής)

Ο κώδικας που αναπτύχθηκε εφαρμόστηκε στην αναφορά των αποτελεσμάτων του Γ' Τριμήνου και Εννεαμήνου 2024 της εταιρείας Helleniq Energy. Στο συγκεκριμένο παράδειγμα, η γρήγορη και η αναλυτική λειτουργία δεν παρουσίασαν σημαντικές διαφορές, καθώς το μέγεθος του κειμένου ήταν σχετικά μικρό. Παρόλα αυτά, ο χρόνος επεξεργασίας παραμένει σημαντικός, δεδομένων των τοπικών δυνατοτήτων του Llama 3 και των υπολογιστικών πόρων του συστήματος που χρησιμοποιήθηκε.

#### Γρήγορη Λειτουργία

## Σύνοψη/Ανάλυση Οικονομικών Δεδομένων

 Γλώσσα Περίληψης:

Ελληνικά

 Επιλογή Λειτουργίας:

Αναλυτική

Γρήγορη

 Γρήγορη Λειτουργία: Επεξεργάζεται μόνο τα πρώτα 3 τμήματα του PDF για μια συνοπτική εκτίμηση. Ιδανική για γρήγορο preview ή μεγάλα αρχεία. (Πιο γρήγορη αλλά λιγότερο αναλυτική)

 Ανέβασε PDF Αναφορά

Drag and drop file here

Limit 200MB per file • PDF

Browse files

Αποτελέσματα Γ' Τριμήνου \_ Εννεαμήνου 2024 - HELLENiQ ENERGY.pdf 0.5MB

Μήκος κειμένου: 14180 χαρακτήρες

Εξήχθη κείμενο (14180 χαρακτήρες)



### Περίληψη Αναφοράς

Ολοκληρώθηκε τμήμα 1/3

Ολοκληρώθηκε τμήμα 2/3

Ολοκληρώθηκε τμήμα 3/3



## Τελική Περίληψη

### Περίληψη:

Η HELLENiQ ENERGY Holdings A.E. δημοσίευσε τα αποτελέσματα Γ' Τριμήνου για το 2024, με ισχυρές λειτουργικές επιδόσεις σε όλες τις δραστηριότητες. Τα Συγκρίσιμα Κέρδη EBITDA ανήλθαν στα €183 εκατ., ενώ τα Συγκρίσιμα Καθαρά Κέρδη στα €49 εκατ. Οι πωλήσεις του τριμήνου αυξήθηκαν κατά 9,7 δισ., ενώ η παραγωγή προϊόντων από τα διυλιστήρια αυξήθηκε κατά 6% στους 3,9 εκατ. τόνους.

Το Δ.Σ. της HELLENiQ ENERGY αποφάσισε τη διανομή προσωρινού μερίσματος €0,20 ανά μετοχή στους μετόχους. Οι κύριες εξελίξεις είναι η υλοποίηση του στρατηγικού προγράμματος του Ομίλου, με πρωτοβουλίες που στοχεύουν στη λειτουργική αριστοποίηση και βελτίωση της κερδοφορίας σε όλες τις δραστηριότητες.

Παράλληλα, η HELLENiQ Renewables λειτουργεί και αναπτύσσει έργα ΑΠΕ συνολικής ισχύος 5,7 GW στην Ελλάδα και διεθνώς. Στον τομέα της Έρευνας και Παραγωγής Υδρογονανθράκων, η επεξεργασία και ερμηνεία των ερευνητικών εργασιών βρίσκεται σε εξέλιξη.

Η HELLENiQ ENERGY συνέχισε να επικεντρωθεί στην ανάπτυξη της θυγατρικής εμπορίας στο εξωτερικό, καθώς και στη βελτίωση των πωλήσεων προϊόντων και υπηρεσιών εκτός καυσίμων. Οι τιμές αργού πετρελαίου (τύπου Brent) υποχώρησαν στα χαμηλότερα επίπεδα από τον Δεκέμβριο 2021, που δημιουργεί περιθώρια διύλισης σε χαμηλότερα επίπεδα.

🕒 Χρόνος περίληψης: 60m 15s (3615.59 sec)



## Βασικά Οικονομικά Μεγέθη

### Συνοπτικά αποτελέσματα:

#### Πωλήσεις:

Όγκοι πωλήσεων: 9,7 δισ. (συγκρίσιμος όγκος πωλήσεων)

Πωλήσεις 9M24: 4,2 εκατ. τόνοι (+8% σε σχέση με το Γ' Τρίμηνο 2016)

#### EBITDA:

EBITDA 9M24: €183 εκατ.

Συσχέτισμα EBITDA: €753 εκατ. (συγκρίσιμα κέρδη EBITDA για το Εννεάμηνο)

#### Καθαρά Κέρδη:

Καθαρά κέρδη 9M24: €284 εκατ.

Συσχέτισμα καθαρών κερδών: €753 εκατ. (συγκρίσιμα καθαρά κέρδη για το Εννεάμηνο)

#### Καθαρός Δανεισμός:

Καθαρός δανεισμός 9M24: €1,77 δισ.

#### Απασχολούμενα Κεφάλαια:

Απασχολουμένων κεφαλαίων 9M24: (δεν αναφέρεται συγκεκριμένο ποσό)

#### Προσωρινό Μέρισμα ανά μετοχή:

Προσωρινό μέρισμα ανά μετοχή: €0,20

#### Επιπτώσεις της τιμής του πετρελαίου:

Τιμές αργού πετρελαίου (τύπου Brent): \$80/βαρέλι (χαμηλότερα σε σχέση με Γ' Τρίμηνο 2023)

Τιμές φυσικού αερίου: +5% σε σχέση με το Γ' Τρίμηνο 2023

Τιμές ηλεκτρικής ενέργειας: +17% σε σχέση με το Γ' Τρίμηνο 2023

Διεθνές περιθώριο αναφοράς συστήματος των διυλιστηρίων: \$3,6/βαρέλι (χαμηλότερα σε σχέση με Γ' Τρίμηνο 2023)

## Στρατηγική

Κύριες εξελίξεις σε κάθε στρατηγικό κλάδο:

**Διύλισης και Πετροχημικών:** Η Εταιρεία βρίσκεται σε εξέλιξη για έργα ενεργειακής αυτονομίας και εξοικονόμησης ενέργειας, ενώ ξεκίνησαν οι εργασίες που αφορούν στην επέκταση της παραγωγικής δυναμικότητας του εργοστασίου παραγωγής πολυπροπυλενίου κατά 60 χιλ. τόνους ετησίως.

**Λιανικής Εμπορίας:** Συνεχίζεται το πρόγραμμα μετεξέλιξης, με εξορθολογισμό και ανάπτυξη του δικτύου λιανικής στην Ελλάδα και στις διεθνείς αγορές, ενίσχυση των ιδιολειτουργούμενων πρατηρίων και των διαφοροποιημένων προϊόντων.

**ΑΠΕ:** Η HELLENiQ Renewables λειτουργεί και αναπτύσσει έργα συνολικής ισχύος 5,7 GW στην Ελλάδα και στη διεθνή αγορά. Έργα ισχύος περίπου 0,4 GW βρίσκονται ήδη σε λειτουργία, ενώ έργα συνολικής δυναμικότητας 0,7 GW υλοποιούνται και αναμένεται να ολοκληρωθούν σταδιακά.

**Έρευνας και Παραγωγής Υδρογονανθράκων:** Σε συνέχεια της ολοκλήρωσης γεωφυσικών καταγραφών σε 5 θαλάσσιες περιοχές, η επεξεργασία και η ερμηνεία των ερευνητικών εργασιών βρίσκεται σε εξέλιξη.

**Γενικό:** Η υλοποίηση του προγράμματος ψηφιακού μετασχηματισμού σε όλες τις δραστηριότητες του Ομίλου συνεχίζεται, με σημαντικά οφέλη που εκτιμώνται, για το 2024, σε περίπου €40 εκατ. ετησίως.

Σε γενικές γραμμές, η Εταιρεία επικεντρώνεται στη βελτίωση της λειτουργικής απόδοσης, στην ασφάλεια προσωπικού και εγκαταστάσεων, στην πρόληψη και στην αποτελεσματικότερη διαχείριση κινδύνων.

 Συνολικός χρόνος επεξεργασίας: 160m 35s (9635.73 sec)

## Σύνοψη/Ανάλυση Οικονομικών Δεδομένων

 Γλώσσα Περίληψης:

Ελληνικά

 Επιλογή Λειτουργίας:

Αναλυτική

Γρήγορη

 Αναλυτική Λειτουργία: Επεξεργάζεται όλο το PDF με μεγαλύτερα τμήματα και περισσότερη λεπτομέρεια. Ιδανική για τελική σύνοψη και πλήρη ανάλυση. (Πιο αργή αλλά πιο ακριβής).

 Ανέβασε PDF Αναφορά

Drag and drop file here

Limit 200MB per file • PDF

Browse files

Αποτελέσματα Γ' Τριμήνου \_ Εννεαμήνου 2024 - HELLENiQ ENERGY.pdf 0.5MB

Μήκος κειμένου: 14180 χαρακτήρες

Εξήχθη κείμενο (14180 χαρακτήρες)

### Περίληψη Αναφοράς

 Διαχωρίστηκαν σε 3 τμήματα (Αναλυτική mode).

Ολοκληρώθηκε τμήμα 1/3

Ολοκληρώθηκε τμήμα 2/3

Ολοκληρώθηκε τμήμα 3/3

### Τελική Περίληψη

Περίληψη στα Ελληνικά:

Η HELLENiQ ENERGY Holdings A.E. ανακοίνωσε τα ενοποιημένα οικονομικά αποτελέσματα Γ' Τριμήνου

2024, με τα Συγκρίσιμα Κέρδη EBITDA να διαμορφώνονται στα €183 εκατ. και τα Συγκρίσιμα Καθαρά Κέρδη στα €49 εκατ.

Η παραγωγή προϊόντων από τα διυλιστήρια αυξήθηκε κατά 6% στους 3,9 εκατ. τόνους, ενώ οι πωλήσεις επίσης αυξήθηκαν. Τα οικονομικά αποτελέσματα επηρεάστηκαν από τη χαμηλή διαθεσιμότητα των διυλιστηρίων.

Στο τμήμα ΑΠΕ, η HELLENIQ Renewables λειτουργεί και αναπτύσσει έργα συνολικής ισχύος 5,7 GW στην Ελλάδα και στη διεθνή αγορά. Τα Κέρδη EBITDA από το τμήμα ΑΠΕ ανήλθαν σε €13 εκατ.

Στα υπόλοιπα τμήματα, η παραγωγή προϊόντων φθάνει σε υψηλά επίπεδα, ενώ οι πωλήσεις αυξήθηκαν. Τα οικονομικά αποτελέσματα επηρεάστηκαν από τα χαμηλότερα περιθώρια διύλισης.

Σε γενικό πλαίσιο, τα αποτελέσματα της εταιρείας είναι θετικά, με αυξήσεις στις πωλήσεις και την κερδοφορία σε όλα τα τμήματα.

🕒 Χρόνος περίληψης: 97m 9s (5829.83 sec)



## Βασικά Οικονομικά Μεγέθη

Ακολουθούν τα συνοπτικά αποτελέσματα:

Πωλήσεις και όγκοι πωλήσεων:

Όγκοι πωλήσεων: 9,7 δισ. (Γ' Τρίμηνο 2024)

Πωλήσεις 9Μ24: +8% σε σχέση με το Γ' Τρίμηνο 2023

Συγκρίσιμα όγκοι πωλήσεων: €0,75 δισ. (Εννεάμηνο 2024)

Κέρδη και EBITDA:

Κέρδη EBITDA: €183 εκατ. (Γ' Τρίμηνο 2024)

Συγκρίσιμα Κέρδη EBITDA: +6% σε σχέση με το Γ' Τρίμηνο 2023

EBITDA: €753 εκατ. (Εννεάμηνο 2024)

Καθαρά Κέρδη και Καθαρός Δανεισμός:

Καθαρά Κέρδη: €284 εκατ. (Εννεάμηνο 2024)

Καθαρός Δανεισμός: €1,77 δισ. (Γ' Τρίμηνο 2024)

Προσωρινό Μίσμα ανά μετοχή:

Προσωρινό μίσμα ανά μετοχή: €0,20

Επιπτώσεις της τιμής του πετρελαίου:

Τιμές αργού πετρελαίου (τύπου Brent): \$80/βαρέλι (Γ' Τρίμηνο 2024)

Τιμές φυσικού αερίου και ηλεκτρικής ενέργειας: +5% και +17% σε σχέση με το Γ' Τρίμηνο 2023

Τιμές δικαιωμάτων εκπομπών CO<sub>2</sub>: -20% σε σχέση με την αντίστοιχη περσινή περίοδο.

## Στρατηγική

Κύριες εξελίξεις σε κάθε στρατηγικό κλάδο:

### 1. Κλάδος Διύλισης και Πετροχημικών:

Υλοποίηση έργων ενεργειακής αυτονομίας και εξοικονόμησης ενέργειας  
Επέκταση της παραγωγικής δυναμικότητας του εργοστασίου παραγωγής πολυπροπυλενίου κατά 60 χιλ. τόνους ετησίως  
Λύσεις ανθρακοποίησης των διεργασιών και βελτίωση του ανθρακικού αποτυπώματος

### 2. Κλάδος Λιανικής Εμπορίας:

Συνεχίζεται το πρόγραμμα μετεξέλιξης με εξορθολογισμό και ανάπτυξη του δικτύου λιανικής στην Ελλάδα και στις διεθνείς αγορές  
Ενίσχυση των ιδιολειτουργούμενων πρατηρίων και των διαφοροποιημένων προϊόντων

### 3. Κλάδος ΑΠΕ:

Λειτουργία έργων συνολικής ισχύος 5,7 GW στην Ελλάδα και στη διεθνή αγορά  
Εγκατάσταση νέων εργοστασίων ΑΠΕ με συνολική δυναμικότητα 0,7 GW

### 4. Κλάδος Έρευνας και Παραγωγής Υδρογονανθράκων:

Επεξεργασία και ερμηνεία των ερευνητικών εργασιών σε 5 θαλάσσιες περιοχές  
Αποφάσεις για ερευνητικές γεωτρήσεις σε πιθανούς στόχους που θα προκύψουν αναμένονται εντός του 2025

### 5. Κλάδος Ψηφιακού Μετασχηματισμού:

Υλοποίηση προγράμματος ψηφιακού μετασχηματισμού σε όλες τις δραστηριότητες του Ομίλου  
Σημαντικά οφέλη που εκτιμώνται για το 2024, σε περίπου €40 εκατ. ετησίως

Γενικά, οι κύριες εξελίξεις του Ομίλου HELLENIQ ENERGY Holdings A.E. είναι η υλοποίηση στρατηγικών προγραμμάτων που στοχεύουν στη βελτίωση της λειτουργικής απόδοσης, την αύξηση της κερδοφορίας και τη μείωση των περιθωρίων διύλισης.

 Συνολικός χρόνος επεξεργασίας: 197m 44s (11864.35 sec)

## (3.6 Chatbot & Prompt Engineering)

---

### (3.6.a Εισαγωγή)

---

Ένα ρομπότ συνομιλίας (chatbot) είναι μια εφαρμογή λογισμικού σχεδιασμένη να μιμείται τη συνομιλία με ανθρώπους. Μπορεί να επικοινωνεί μέσω κειμένου ή φωνής και να απαντά σε ερωτήσεις των χρηστών με φυσικό τρόπο.

Τα απλά chatbots λειτουργούν βάσει προκαθορισμένων κανόνων και σεναρίων, παρέχοντας περιορισμένες και προβλέψιμες απαντήσεις σε συγκεκριμένες λέξεις ή φράσεις, χωρίς δυνατότητα μάθησης από τις συνομιλίες. Αντίθετα, ένα AI chatbot, όπως το ChatGPT ή το LLaMA 3, χρησιμοποιεί μηχανική μάθηση και επεξεργασία φυσικής γλώσσας (NLP), κατανοεί σύνθετες ερωτήσεις και συμφραζόμενα, προσαρμόζει τις απαντήσεις του στον χρήστη και μπορεί να βελτιώνεται συνεχώς με νέα δεδομένα, προσφέροντας πιο φυσική, ευέλικτη και «έξυπνη» συνομιλία. Στο πεδίο των χρηματοοικονομικών εφαρμογών, αυτά τα chatbots μπορούν να παρέχουν αναλύσεις για μετοχές, προβλέψεις αγοράς ή εξατομικευμένες επενδυτικές συμβουλές.

Στον χώρο της τεχνητής νοημοσύνης και των γενετικών μοντέλων, ο όρος prompt αναφέρεται στο ερέθισμα κειμένου ή είσοδο που δίνει ο χρήστης για να καθοδηγήσει την έξοδο του μοντέλου. Το prompt μπορεί να είναι μια απλή ερώτηση, μια σύνθετη περιγραφή ή ένα σύνολο οδηγιών. Ένα αποτελεσματικό prompt καθοδηγεί τη συμπεριφορά του μοντέλου, ενώ πρόσθετα στοιχεία, όπως δεδομένα εισόδου ή παραδείγματα, μπορούν να βελτιώσουν σημαντικά την ποιότητα της απάντησης.

Η μηχανική προτροπών (prompt engineering) αποτελεί μια ταχέως αναπτυσσόμενη πρακτική, με στόχο τη βελτιστοποίηση της αλληλεπίδρασης μεταξύ χρηστών και LLM και την επίτευξη υψηλής ποιότητας αποτελεσμάτων [14]. Δεν αφορά μόνο τη σύνταξη ερωτήσεων, αλλά και τη στρατηγική χρήση οδηγιών, συμφραζομένων και παραδειγμάτων, λαμβάνοντας υπόψη τις δυνατότητες και τους περιορισμούς του μοντέλου. Στην πράξη, η διαδικασία είναι επαναληπτική, διερευνητική και πειραματική, με συνεχή βελτιστοποίηση των prompts για σταθερά και αξιόπιστα αποτελέσματα.

Οι κυριότερες τεχνικές prompt engineering παρουσιάζονται στον Πίνακα 3.9.

Στο πλαίσιο της χρηματιστηριακής ανάλυσης, η σωστή χρήση prompts και τεχνικών prompt engineering επιτρέπει σε LLM να επεξεργάζονται δεδομένα μετοχών, να εξάγουν σημαντικά συμπεράσματα και να παρέχουν ακριβέστερες προβλέψεις, καθιστώντας τα πολύτιμα εργαλεία για επενδυτικές αποφάσεις.

Πίνακας 3.9: Τεχνικές Prompt Engineering

Τεχνική	Περιγραφή / Σκοπός	Πλεονεκτήματα	Ενδεικτικές Εφαρμογές
<b>Zero-Shot Prompting</b>	Το μοντέλο εκτελεί μια εργασία χωρίς κανένα παράδειγμα ή επιπλέον καθοδήγηση	Πολύ απλό και άμεσο	Αυτόματες μεταφράσεις
		Δεν απαιτεί προετοιμασία παραδειγμάτων	Γενικές απαντήσεις ή περιλήψεις
		Κατάλληλο για γενικές εργασίες	Chatbots
<b>Few-Shot Prompting</b>	Παροχή λίγων παραδειγμάτων για να «μάθει» το μοντέλο τον επιθυμητό τρόπο απάντησης	Αυξάνει την ακρίβεια και συνέπεια	Δημιουργία περιεχομένου με συγκεκριμένο ύφος.
		Ενσωματώνει ύφος, δομή ή πλαίσιο	Επεξεργασία φυσικής γλώσσας με προσαρμογή στον τομέα
<b>Chain of Thought (CoT)</b>	Καθοδηγεί το μοντέλο να «σκεφτεί φωναχτά», εκφράζοντας τα ενδιάμεσα βήματα λογισμού	Ενισχύει τη λογική συνέπεια	Μαθηματικά προβλήματα
		Αυξάνει την ακρίβεια σε σύνθετα προβλήματα	Λογικοί συλλογισμοί
			Αναλυτικές ερωτήσεις
<b>Self-Consistency</b>	Δημιουργία πολλών διαφορετικών λογικών ακολουθιών και επιλογή της πιο συνεπούς απάντησης	Μειώνει τα λάθη λογισμού	Επιστημονική επίλυση προβλημάτων
		Παράγει πιο αξιόπιστες απαντήσεις	Επαληθεύσεις δεδομένων
<b>ReAct (Reason + Act)</b>	Συνδυάζει σκέψη και δράση — το μοντέλο «σκέφτεται» και στη συνέχεια ενεργεί (π.χ. κάνει αναζητήσεις)	Δυνατότητα ενεργού λήψης αποφάσεων	Ευφυείς πράκτορες (AI agents)
		Ικανότητα αλληλεπίδρασης με περιβάλλον	Συνομιλητικά συστήματα με εξωτερική πρόσβαση (π.χ. web)
<b>Prompt Chaining</b>	Διάσπαση μιας πολύπλοκης εργασίας σε διαδοχικά prompts που συνεργάζονται μεταξύ τους	Αυξάνει την ακρίβεια και σαφήνεια	Δημιουργία περιεχομένου
		Επιτρέπει πολυσταδιακή λογική	Ανάλυση δεδομένων
			Πολυβήματες διεργασίες
<b>Automatic Prompt Optimization (APO)</b>	Αυτόματη βελτιστοποίηση prompts μέσω ανατροφοδότησης ή αλγοριθμικής διαδικασίας	Εξοικονομεί χρόνο	Βιομηχανικές εφαρμογές
		Προσαρμόζεται δυναμικά	Chatbots μεγάλης κλίμακας
		Μαθαίνει από τα αποτελέσματα	Παραγωγή διαφημιστικού περιεχομένου

### (3.6.6 Περιγραφή Εφαρμογής Chatbot)

Αναπτύχθηκε μια εφαρμογή chatbot χρησιμοποιώντας το τοπικό μοντέλο LLaMA 3 σε συνδυασμό με τεχνικές prompt engineering, με στόχο την παραγωγή έξυπνων και ενημερωμένων απαντήσεων σε ερωτήματα χρηστών, ειδικά σε θέματα χρηματιστηρίου. Η εφαρμογή ακολουθεί τα εξής βασικά βήματα:

1. Ο χρήστης πληκτρολογεί το ερώτημά του σε διεπαφή τύπου streamlit.
2. Η εφαρμογή πραγματοποιεί αναζήτηση στο διαδίκτυο μέσω της μηχανής DuckDuckGo, παίρνοντας τα τρία πρώτα αποτελέσματα. Τα αποτελέσματα περιλαμβάνουν τίτλο, σύντομο απόσπασμα και URL, τα οποία θα χρησιμοποιηθούν για την ενίσχυση του prompt.
3. Με βάση τα αποτελέσματα της αναζήτησης και την ερώτηση του χρήστη, δημιουργείται ένα πλήρες prompt που καθοδηγεί το μοντέλο να απαντήσει με ακρίβεια και να παραθέσει τις πηγές (URLs).
4. Το prompt αποστέλλεται στο τοπικό μοντέλο LLaMA 3, μέσω του API της Ollama. Το μοντέλο επιστρέφει την απάντηση και τον χρόνο εκτέλεσης, ενώ ο κώδικας διαχειρίζεται πιθανά σφάλματα σύνδεσης ή προβλήματα στην απόκριση JSON.
5. Η διεπαφή streamlit εμφανίζει την απάντηση του μοντέλου μαζί με τα URLs που χρησιμοποιήθηκαν και τον χρόνο επεξεργασίας.

Αρχικά γίνεται αίτημα στο τοπικό μοντέλο Llama 3 το οποίο επιστρέφει την απάντηση του μοντέλου και τον χρόνο εκτέλεσης και διαχειρίζεται πιθανά σφάλματα.

```
def llama_3_ollama(prompt, model="Llama 3", timeout=1000):
    start_time = time.time()
    try:
        response = requests.post(
            "http://localhost:11434/api/generate",
            json={"model": model, "prompt": prompt, "stream": False},
            timeout=timeout
        )
        response.raise_for_status()
        data = response.json()
        answer = data.get("response", "").strip()
    except requests.exceptions.RequestException as e:
        answer = f"⚠️ Connection error: {e}"
    except ValueError:
        answer = "⚠️ Invalid JSON response from Ollama."
    elapsed = time.time() - start_time
    return answer, elapsed
```

Στη συνέχεια εκτελείται αναζήτηση στο web μέσω DuckDuckGo Search (μηχανή αναζήτησης) για το ερώτημα του χρήστη, παίρνει τα 3 πρώτα αποτελέσματα και επιστρέφει τίτλο, μικρό απόσπασμα και URL.

```
def search_web(query, num_results=3):
```

```

try:
    with DDGS() as ddgs:
        results = list(ddgs.text(query, max_results=num_results))
    if not results:
        return "No search results found."
    snippets = "\n\n".join([
        f>Title: {r.get('title','No title')}\nSnippet: {r.get('body','')[:300]}...\nURL: {r.get('href','')}"
        for r in results
    ])
    return snippets
except Exception as e:
    return f"⚠ Search error: {e}"

```

Η συνάρτηση `ask_stateless` συνδυάζει την αναζήτηση και την κλήση του μοντέλου, δημιουργώντας ένα prompt.

```

def ask_stateless(question, model="Llama 3"):
    search_results = search_web(question)
    prompt = (
        "You are a helpful AI assistant.\n"
        f"Use the following web search results to answer the user's question:\n{search_results}\n\n"
        f"User's question: {question}\n"
        "Answer concisely and cite URLs when possible."
    )
    return Llama 3_ollama(prompt, model=model)

```

Τέλος, η διεπαφή `streamlit` επιτρέπει στον χρήστη να εισάγει το ερώτημά του, να επιλέξει μοντέλο και να λάβει γρήγορη απάντηση με αναφορά στις πηγές και το χρόνο επεξεργασίας.

```

st.set_page_config(page_title="🌐 LLaMA 3 Web Chat", layout="wide")
st.title("🌐 LLaMA 3 Web Chat")

question = st.text_area("Ask a question:", height=300)
model = st.selectbox("Choose a local Ollama model:", ["Llama 3", "Llama 3.1", "mistral", "gemma2"])

if st.button("Ask") and question.strip():
    with st.spinner("🔍 Searching and thinking..."):
        answer, elapsed = ask_stateless(question, model=model)
    st.markdown(f"**Answer:** {answer}")
    st.markdown(f"_Took {elapsed:.2f} seconds_")

```

### (3.6.γ Παραδείγματα Εφαρμογής Chatbot & Prompt Engineering)

---

Η εφαρμογή αξιοποιήθηκε σε συνδυασμό με διάφορες τεχνικές prompt engineering, προσαρμοσμένες για την επεξεργασία ερωτημάτων και την αναζήτηση απαντήσεων σε θέματα χρηματιστηρίου. Για τη συζήτηση των αποτελεσμάτων ανατρέξτε στην παράγραφο 4.4.

#### Zero-Shot Prompting

---

Το Zero-Shot Prompting είναι μια τεχνική στην επεξεργασία φυσικής γλώσσας που επιτρέπει σε προ-εκπαιδευμένα μοντέλα γλώσσας να εκτελούν νέες εργασίες χωρίς προηγούμενη εκπαίδευση ή παραδείγματα για την ίδια εργασία.

#### Ερώτηση

Να δοθούν πληροφορίες για το χρηματιστηριακό σύμβολο AAPL (Apple Inc.), όπως όνομα εταιρείας, τομέας και πρόσφατη χρηματιστηριακή απόδοση.

#### Prompt

```
You are a financial assistant that provides concise, factual summaries about stock tickers.
Task: For a given stock ticker, return the company name, the sector/industry, and a short summary of
recent stock performance.

Ticker: AAPL

Answer format:
- Company Name:
- Sector:
- Recent Performance:
- Notes:
```

#### Απάντηση Μοντέλου

```
Answer: Based on the provided web search results, here is the answer for the given stock ticker:

Company Name: Apple Inc.
Sector: Technology
Recent Performance: As of Yahoo Finance's latest update, AAPL has shown a positive trend with a 1-
day return of +0.43% and a 5-day return of +2.33%. (Source: https://finance.yahoo.com/)
Notes: No additional information available from the provided search results.

Please note that the recent performance may have changed since the last update on Yahoo Finance,
which is subject to change and may not reflect the current market situation.

Took 259.36 seconds
```

Το Few-Shot Prompting είναι μια τεχνική όπου παρέχονται στο μοντέλο λίγα παραδείγματα (συνήθως 2–5) για να κατανοήσει τη δομή και τον τρόπο απάντησης που απαιτείται. Αυτή η μέθοδος επιτρέπει στο μοντέλο να παράγει πιο ακριβείς, συνεπείς και οργανωμένες απαντήσεις χωρίς εκπαίδευση σε ολόκληρο dataset. Στο πλαίσιο χρηματιστηριακής ανάλυσης, χρησιμοποιείται για να παρέχει συνοπτικές, τεκμηριωμένες πληροφορίες για διάφορα χρηματιστηριακά σύμβολα (stock tickers).

### Ερώτηση

Να δοθούν πληροφορίες για το χρηματιστηριακό σύμβολο MSFT (Microsoft Corp.), χρησιμοποιώντας προηγούμενα παραδείγματα για να καθοδηγηθεί το μοντέλο στην σωστή μορφή και περιεχόμενο της απάντησης δίνοντάς πληροφορίες για όνομα εταιρείας, τομέας, πρόσφατη χρηματιστηριακή απόδοση και σημειώσεις.

### Prompt

You are a financial assistant that provides concise, factual summaries about stock tickers. Use the following examples as a guide and keep your answers in the same format.

#### Examples:

Ticker: AAPL

- Company Name: Apple Inc.
- Sector: Technology (Consumer Electronics)
- Recent Performance: Stable growth with steady demand for iPhone and strong services revenue.
- Notes: Large-cap stock often used as a benchmark for tech performance.

Ticker: TSLA

- Company Name: Tesla, Inc.
- Sector: Consumer Cyclical (Automotive)
- Recent Performance: Moderate volatility with fluctuations tied to production numbers and EV market trends.
- Notes: Strong growth potential but sensitive to interest rate changes and competition.

Ticker: NVDA

- Company Name: NVIDIA Corporation
- Sector: Technology (Semiconductors)
- Recent Performance: Rapid growth driven by demand for GPUs and AI hardware.
- Notes: Considered a leading company in AI infrastructure.

Now complete the following in the same format.

## Απάντηση Μοντέλου

Answer: Here is the answer:

Ticker: MSFT

Company Name: Microsoft Corporation

Sector: Technology (Software)

Recent Performance: Consistent growth driven by cloud computing services, Azure, and Office 365.

Notes: Large-cap stock with a diverse portfolio of products and services.

Source: <https://finance.yahoo.com/quote/MSFT/profile>

(Note: The recent performance and notes are based on general market trends and may not reflect the current exact situation. For up-to-date information, please visit the provided URL or other reliable financial sources.)

*Took 328.71 seconds*

Το Chain-of-Thought (CoT) είναι μια τεχνική όπου το μοντέλο καθοδηγείται να σκέφτεται βήμα-βήμα προτού δώσει την τελική απάντηση. Αυτή η μέθοδος ενισχύει την ικανότητα του μοντέλου να επεξεργάζεται σύνθετα ερωτήματα, αναλύει ενδιάμεσα βήματα και παρέχει τεκμηριωμένες και συνεπείς απαντήσεις. Στο χρηματιστηριακό πλαίσιο, το CoT χρησιμοποιείται για να αξιολογήσει επενδυτικές αποφάσεις, λαμβάνοντας υπόψη πολλούς παράγοντες όπως τάσεις αγοράς, οικονομικά δεδομένα και κινδύνους.

### Ερώτηση

Να αξιολογηθεί αν ένας επενδυτής θα πρέπει να εξετάσει την αγορά της μετοχής NVIDIA (NVDA), χρησιμοποιώντας ενδιάμεσα βήματα σκέψης για πιο τεκμηριωμένη ανάλυση. Τα βήματα σκέψης περιλαμβάνουν:

1. Κατανόηση του αντικειμένου της NVIDIA και των κύριων επιχειρηματικών τομέων
2. Αξιολόγηση της πρόσφατης απόδοσης της αγοράς και του κλάδου
3. Ανάλυση πρόσφατων τάσεων της εταιρείας (π.χ. ζήτηση για AI, πωλήσεις, ανάπτυξη κερδών)
4. Εκτίμηση πιθανών κινδύνων ή μεταβλητότητας
5. Σύνθεση των παραπάνω σε μια σύντομη και σαφή επενδυτική άποψη

### Prompt

You are a financial analyst.

Question: Should an investor consider buying stock in NVIDIA (NVDA) at this time, based on recent performance trends?

Let's think step by step.

1. Identify what NVIDIA does and its main business areas.
2. Consider the recent market and sector performance.
3. Evaluate any recent company trends (e.g., AI demand, GPU sales, earnings growth).
4. Consider potential risks or volatility.
5. Then, summarize with a concise investment view.

### Απάντηση Μοντέλου

Answer: I'd be happy to help you make an informed decision about NVIDIA (NVDA) stock.

1. What does NVIDIA do? NVIDIA is a leader in the semiconductor industry, specializing in graphics processing units (GPUs) and high-performance computing hardware, with a focus on artificial intelligence (AI), gaming, and datacenter applications.

2. Recent market and sector performance The semiconductor industry has been experiencing volatility recently, with concerns over trade tensions and supply chain disruptions (Source: Stockchase). However, AI-related spending has remained strong, driving demand for NVIDIA's products.

3. Recent company trends Earnings growth has slowed down this year, but NVIDIA's dominant market position and increased AI investment remain positive factors (Sources: Fool.com, Morningstar).

4. Potential risks or volatility Underlying concerns about valuation and margin sustainability exist, as some experts caution against recent volatility and external factors like trade tensions.

Summary and Investment View Based on the analysis, I would recommend taking a cautious approach when considering buying NVIDIA stock at this time. While AI demand and government spending on data centers remain positive trends, the company's earnings growth slowdown and sector volatility warrant careful consideration. It may be wise to wait for the quarterly financial report to restore investor confidence before making a decision (Source: Fool.com).

Please note that this analysis is based on publicly available information and should not be considered personalized investment advice.

*Took 486.89 seconds*

## Self-Consistency

Η τεχνική Self-Consistency ζητά από το μοντέλο να παράγει πολλές ανεξάρτητες λογικές διαδρομές (reasoning paths) για το ίδιο ερώτημα. Κάθε διαδρομή σκέψης αξιολογεί διαφορετικούς παράγοντες ή σενάρια. Η τεχνική αυξάνει την αξιοπιστία της απάντησης, αφού οι αποκλίσεις ανάμεσα στις διαδρομές επιτρέπουν την επιλογή της πιο συνεπούς ή λογικής απάντησης. Στο χρηματιστηριακό πλαίσιο, χρησιμοποιείται για να εξετάσει πολλαπλές απόψεις για μια επενδυτική απόφαση, μειώνοντας τα μεροληπτικά ή μονοδιάστατα αποτελέσματα.

### Ερώτηση

Να αξιολογηθεί αν ένας επενδυτής θα πρέπει να εξετάσει την αγορά της μετοχής Tesla (TSLA) με βάση πρόσφατες επιδόσεις, τάσεις της αγοράς και κινδύνους της βιομηχανίας, χρησιμοποιώντας τρία ανεξάρτητα μονοπάτια σκέψης για μεγαλύτερη εγκυρότητα.

Βήματα κάθε διαδρομής:

1. Ανάλυση των βασικών επιχειρηματικών στοιχείων της Tesla
2. Αξιολόγηση των πρόσφατων τάσεων της αγοράς και των ειδήσεων
3. Συζήτηση πιθανών κινδύνων και ευκαιριών
4. Σύντομη επενδυτική σύσταση

### Prompt

Should an investor consider buying Tesla (TSLA) stock at this time, given recent company performance, market trends, and industry risks?

Generate 3 independent reasoning paths. Each path should:

1. Analyze Tesla's business fundamentals.
2. Consider recent market trends and news.
3. Discuss potential risks and opportunities.
4. Conclude with a short investment recommendation.

### Απάντηση Μοντέλου

Answer: Here are three independent reasoning paths for considering buying Tesla (TSLA) stock:

Path 1: Fundamentals, Trends, and Risks

Tesla's recent quarterly results show consistent growth in revenue and profitability. The company's focus on electric vehicle (EV) production and autonomous driving technology positions it well for long-term success. Recent market trends favor TSLA, with the S&P 500 index performing well despite global economic concerns.

However, industry risks include increasing competition from established automakers and new entrants in the EV space. Regulatory scrutiny of the auto industry's carbon emissions and potential changes to

government incentives could impact demand.

Recommendation: Consider buying TSLA for its strong fundamentals and growth prospects, but maintain a cautious approach due to industry risks (Source: [1])

#### Path 2: Cryptocurrency and Market Volatility

The recent surge in cryptocurrency adoption and market volatility may have implications for the automotive sector. As cryptocurrencies like Bitcoin and Ethereum gain traction, investors are seeking diversification opportunities. Tesla's early mover advantage in EV production and autonomous driving technology could attract crypto-investors seeking exposure to the sector.

However, Trump's memecoins, which raise ethical concerns, may lead to increased regulatory scrutiny of the entire cryptocurrency market, including potential effects on the automotive industry.

Recommendation: Consider buying TSLA as a hedge against potential market volatility and uncertainty in the cryptocurrency space (Source: [2])

#### Path 3: Regulatory Environment

The SEC's efforts under Gary Gensler's leadership may lead to increased regulatory scrutiny of the entire crypto-market, including memecoins like Trump's. This could have implications for Tesla's future expansion plans and partnerships.

On the positive side, a clearer regulatory framework might encourage more mainstream investment in the automotive sector, benefiting companies like TSLA.

Recommendation: Consider buying TSLA as a way to play the potential long-term growth of the electric vehicle market, while being mindful of the regulatory environment (Source: [3])

Please note that these paths are independent and not necessarily aligned with each other. It's essential for investors to conduct their own research and consider multiple factors before making an investment decision.

*Took 833.25 seconds*

## ReAct (Reason + Act)

---

Η τεχνική ReAct συνδυάζει δύο βασικά στοιχεία:

- Σκέψη (Reason): το μοντέλο σκέφτεται βήμα-βήμα πώς θα απαντήσει σε ένα ερώτημα.
- Ενέργεια (Action): το μοντέλο εκτελεί ενέργειες, όπως αναζήτηση στο διαδίκτυο ή πρόσβαση σε δεδομένα.

Με αυτόν τον τρόπο, η απάντηση γίνεται πιο ακριβής και ενημερωμένη, ειδικά όταν απαιτείται πρόσφατη ή δυναμική πληροφορία, όπως οι χρηματιστηριακές τιμές.

### Ερώτηση

Να εξεταστεί η πρόσφατη τάση της μετοχής Apple (AAPL) και να παρασχεθεί μια σύντομη, τεκμηριωμένη ανάλυση. Βήματα σκέψης και ενέργειας:

- Σκέψη (Reason): Σκέψη για το πώς θα αναλυθεί η μετοχή.
- Ενέργεια (Action): Λήψη δεδομένων από πηγές όπως το Yahoo Finance.
- Παρατήρηση (Observation): Αναφορά των αποτελεσμάτων.
- Δικαιολόγηση (Reasoning): Ανάλυση των δεδομένων.
- Απάντηση (Answer): Τελική σύντομη σύνοψη για τον επενδυτή.

### Prompt

You are a financial assistant that can reason and take actions.

Question:

Check the recent stock price trend for Apple (AAPL) and provide a brief summary.

Use the following format:

- Reason: Think through how to analyze the stock.
- Action: Specify what you would do (e.g., fetch data from an API).
- Observation: Report the result of the action.
- Reasoning: Analyze the observation.
- Answer: Give the final summary.

### Απάντηση Μοντέλου

Answer: Here's my response:

Reason: To analyze the recent stock price trend for Apple (AAPL), I'll need to access its historical stock data.

Action: I'll fetch the latest stock quote, history, and news from Yahoo Finance using their API:  
<https://finance.yahoo.com/quote/AAPL/>.

Observation: According to Yahoo Finance, as of March 15th, the closing price of AAPL was 184.66 and a one-year low of \$133.49.

Reasoning: A quick glance at the data suggests that AAPL's stock price has been trending upward over the past year, with some fluctuations along the way. The recent closing price is close to its one-year high, indicating potential growth momentum.

Answer: Based on my analysis, Apple (AAPL) appears to be in an upward trend, suggesting a positive outlook for investors. However, as with any investment, it's essential to consider multiple factors before making a decision. For further insights, I recommend exploring additional data and metrics using a free stock screener like StockAnalysis: <https://stockanalysis.com/stocks/screener/>.

*Took 372.63 seconds*

## Prompt Chaining

Η τεχνική Prompt Chaining χωρίζει μια πολύπλοκη εργασία σε διαδοχικά στάδια:

1. Κάθε στάδιο έχει ξεχωριστό prompt και στόχο.
2. Το αποτέλεσμα κάθε σταδίου χρησιμοποιείται ως είσοδος για το επόμενο.
3. Επιτρέπει πιο οργανωμένες και αναλυτικές απαντήσεις, μειώνοντας την πιθανότητα παραλείψεων ή σφαλμάτων.

Στη χρηματοοικονομική ανάλυση, χρησιμοποιείται για σταδιακή αξιολόγηση εταιρειών, ξεκινώντας από οικονομικά στοιχεία, συνεχίζοντας με τάσεις αγοράς και καταλήγοντας σε επενδυτικές συστάσεις.

### Ερώτηση

Να παρασχεθεί μια ολοκληρωμένη και δομημένη επενδυτική αξιολόγηση για τη μετοχή Tesla (TSLA), ακολουθώντας τρία διαδοχικά βήματα:

1. Συνοπτική παρουσίαση των οικονομικών αποτελεσμάτων
2. Ανάλυση τάσεων αγοράς με βάση τη συνοπτική αναφορά
3. Σύντομη επενδυτική σύσταση με βάση τα προηγούμενα αποτελέσματα

### Prompt

**Step 1 – Financial Summary:**

You are a financial analyst.

Task: Summarize Tesla's (TSLA) most recent quarterly earnings and overall financial performance in 3-4 concise sentences. Output only the summary.

**Step 2 – Market Trend Analysis:**

You are a financial analyst.

Task: Analyze Tesla's market trends based on the financial summary. Consider EV sector growth, competitor activity, and potential risks. Output a concise market trend analysis.

Financial Summary:

"Tesla reported strong quarterly earnings with revenue of \$25.3B, up 20% YoY. Profit margins improved due to higher vehicle deliveries and energy product growth. Cash flow remains strong, supporting continued expansion. Overall, Tesla's financial fundamentals are solid."

**Step 3 – Investment Recommendation:**

You are a financial analyst.

Task: Based on the financial summary and market trend analysis, provide a short investment recommendation for Tesla stock. Include a brief rationale.

Financial Summary:

"Tesla reported strong quarterly earnings with revenue of \$25.3B, up 20% YoY. Profit margins improved due to higher vehicle deliveries and energy product growth. Cash flow remains strong, supporting continued expansion. Overall, Tesla's financial fundamentals are solid."

Market Trend Analysis: "Tesla benefits from rapid EV sector growth and increasing global demand for electric vehicles. Competitors are expanding, but Tesla maintains a technology and brand advantage. Risks include regulatory changes, production bottlenecks, and high stock valuation."

## Απάντηση Μοντέλου

Answer: Here are the answers:

Step 1 - Financial Summary: Tesla reported strong quarterly earnings with revenue of \$25.3B, up 20% YoY.

Profit margins improved due to higher vehicle deliveries and energy product growth. Cash flow remains strong, supporting continued expansion. Overall, Tesla's financial fundamentals are solid. (Source: [investing.com/equities/tesla-motors](https://investing.com/equities/tesla-motors))

Step 2 - Market Trend Analysis: Tesla benefits from rapid EV sector growth and increasing global demand for electric vehicles. Competitors are expanding, but Tesla maintains a technology and brand advantage.

Risks include regulatory changes, production bottlenecks, and high stock valuation. (Source: [cnbc.com/2025/01/29/tesla-tsla-2024-q4-earnings.html](https://cnbc.com/2025/01/29/tesla-tsla-2024-q4-earnings.html))

Step 3 - Investment Recommendation: I recommend investing in Tesla stock (TSLA) due to its strong financial fundamentals and dominant position in the EV sector. While there are risks, I believe Tesla's technology and brand advantages will continue to drive growth. Consider a long-term investment strategy.

*Took 546.73 seconds*

## Automatic Prompt Optimization (APO)

Η τεχνική APO επιτρέπει σε ένα LLM να βελτιστοποιεί το ίδιο το prompt που του δίνεται, προτείνοντας εναλλακτικούς τρόπους διατύπωσης για καλύτερη ποιότητα απάντησης.

- Το μοντέλο αξιολογεί το αρχικό prompt
- Προτείνει 2–3 εναλλακτικές διατυπώσεις
- Παράγει απαντήσεις για κάθε πρόταση
- Επιλέγει ποια διατύπωση είναι η πιο αποτελεσματική

Αυτή η τεχνική είναι χρήσιμη όταν η ακρίβεια, η σαφήνεια και η συντομία της απάντησης είναι κρίσιμες, όπως σε χρηματοοικονομικές αναλύσεις ή επενδυτικές συστάσεις.

### Ερώτηση

Να παραχθεί μια συνοπτική και κατανοητή ανάλυση των τριμηνιαίων αποτελεσμάτων της Tesla (TSLA), ενώ ταυτόχρονα να βελτιωθεί η ποιότητα του prompt ώστε να είναι πιο σαφές και σχετικό για επενδυτές.

### Prompt

You are a financial analyst assistant.

Task: Generate a concise, clear, and investor-relevant summary of Tesla's (TSLA) most recent quarterly earnings. Include revenue, profit margins, cash flow, growth trends, and key insights for investors.

I want you to suggest **\*\*3** alternative ways to phrase this prompt**\*\*** that could potentially improve the quality of the summary. Then, for each suggested prompt, generate the corresponding summary. Finally, recommend which prompt is the most effective and why.

### Απάντηση Μοντέλου

Answer: Here are three alternative ways to phrase the original prompt:

1. Reframe: "Provide a concise analysis of Tesla's (TSLA) latest quarterly earnings report, highlighting revenue, profit margins, cash flow, growth trends, and key insights for investors."
2. Expand: "Summarize Tesla's (TSLA) most recent quarterly earnings in 150-200 words, focusing on the company's financial performance, including revenue, gross margin, operating income, net income, cash flow, and year-over-year growth rates."
3. Narrow: "Deliver a brief summary of Tesla's (TSLA) Q4 2024 earnings report, emphasizing key metrics such as revenue, profit margins, and cash flow, as well as any notable trends or insights for investors."

Based on the original prompt, here is the summary:

According to CNBC's article, Tesla missed on earnings and revenue for its fourth quarter. In 2022, Tesla

reported a revenue of \$81 billion (Macrotrends). The company has been struggling with production costs and supply chain issues, affecting its profit margins. Cash flow has also been impacted, with Tesla generating negative free cash flow in the past few years.

Alternative summaries:

Reframe:

Based on CNBC's report, Tesla's Q4 earnings fell short of expectations. Revenue came in at \$81 billion in 2022 (Macrotrends), but profit margins were affected by production costs and supply chain issues. Cash flow has been a concern, with negative free cash flow reported in recent years. Key takeaways for investors include the company's continued focus on ramping up production and improving efficiency.

Expand:

In its Q4 2024 earnings report, Tesla fell short of expectations, reporting revenue of 2.3 billion. Cash flow has been a concern, with negative free cash flow reported in recent years. Year-over-year growth rates have slowed, highlighting the need for Tesla to improve its financial performance.

Narrow:

Tesla's Q4 2024 earnings report saw revenue come in at \$81 billion (Macrotrends). Profit margins were impacted by production costs and supply chain issues, resulting in a gross margin of 20.5%. Cash flow has been a concern, with negative free cash flow reported. Key takeaways for investors include the company's continued focus on ramping up production and improving efficiency.

Recommendation:

The Reframe prompt is the most effective because it provides a clear summary while also highlighting key points and insights for investors. The rephrased prompt focuses on delivering a concise and investor-relevant analysis, which aligns with the original prompt's intent.

*Took 695.06 seconds*

### (3.7 Red Flagging για Διαχείριση Κινδύνων)

---

Στον χώρο των χρηματιστηριακών αγορών, ο όρος *red flagging* αναφέρεται στην έγκαιρη αναγνώριση προειδοποιητικών ενδείξεων ή ύποπτων μοτίβων που μπορεί να υποδηλώνουν αυξημένο επενδυτικό κίνδυνο, προβλήματα στη χρηματοοικονομική υγεία μιας εταιρείας ή πιθανές αποκλίσεις από ορθές εταιρικές πρακτικές. Η διαδικασία αυτή αποτελεί αναπόσπαστο μέρος της διαχείρισης κινδύνων (*risk assessment*), καθώς επιτρέπει στους επενδυτές και στους αναλυτές να εντοπίζουν ενδεχόμενες απειλές πριν αυτές εξελιχθούν σε σοβαρά οικονομικά ή λειτουργικά ζητήματα. Η έγκαιρη ανίχνευση τέτοιων σημάτων συμβάλλει στη λήψη πιο ενημερωμένων και συνετών αποφάσεων, στην προστασία του επενδυτικού κεφαλαίου και στη διατήρηση διαφάνειας στις εταιρικές δραστηριότητες.

Η έννοια του *red flagging* λειτουργεί ως αρχικός μηχανισμός εκκίνησης της διαδικασίας αξιολόγησης κινδύνου. Αρχικά, τα *red flags* βοηθούν στον πρώιμο εντοπισμό περιοχών που ενδέχεται να κρύβουν αυξημένη επικινδυνότητα. Παραδείγματα τέτοιων ενδείξεων μπορεί να είναι ξαφνικές μεταβολές οικονομικών δεικτών, μη αναμενόμενη αύξηση χρέους, απότομη πτώση των περιθωρίων κέρδους, συχνές αλλαγές στη διοικητική ομάδα ή καθυστερήσεις στη δημοσίευση οικονομικών καταστάσεων. Σε αυτό το στάδιο, τα σήματα δεν αποτελούν επιβεβαίωση προβλήματος, αλλά λειτουργούν ως κατευθυντήριες ενδείξεις για περαιτέρω διερεύνηση.

Αφού εντοπιστούν οι πρώτες ενδείξεις, ακολουθεί η κατηγοριοποίηση των κινδύνων. Κάθε *red flag* συνδέεται συνήθως με συγκεκριμένο τύπο κινδύνου, όπως χρηματοοικονομικό, λειτουργικό, κανονιστικό ή κίνδυνο φήμης. Η κατηγοριοποίηση αυτή επιτρέπει την ιεράρχηση και τη συστηματική διερεύνηση των παραγόντων που θα μπορούσαν να επηρεάσουν την αξία της μετοχής ή τη σταθερότητα της εταιρείας. Το επόμενο στάδιο είναι η αξιολόγηση της σημαντικότητας των κινδύνων, όπου εκτιμάται τόσο η πιθανότητα εμφάνισής τους όσο και η πιθανή επίπτωσή τους. Για παράδειγμα, μια απότομη αύξηση του δανεισμού σε συνδυασμό με πτώση των ταμειακών ροών αποτελεί σαφές σήμα υψηλής σοβαρότητας που μπορεί να υποδεικνύει μελλοντική αφερεγγυότητα.

Στη συνέχεια, τα *red flags* ενσωματώνονται στη διαδικασία λήψης αποφάσεων και καθορισμού μέτρων μετριασμού κινδύνου. Ανάλογα με τη σοβαρότητα των ευρημάτων, ο επενδυτής μπορεί να αποφασίσει να πραγματοποιήσει πρόσθετο έλεγχο δέουσας επιμέλειας, να αναπροσαρμόσει το χαρτοφυλάκιό του, να καθυστερήσει μια επενδυτική τοποθέτηση ή να ενισχύσει την παρακολούθηση συγκεκριμένων χρηματοοικονομικών δεικτών. Το *red flagging* επιτρέπει, έτσι, πιο στοχευμένη και τεκμηριωμένη διαχείριση κινδύνου, ενισχύοντας τη στρατηγική λήψης αποφάσεων.

Τέλος, η διαδικασία δεν είναι στατική, καθώς οι χρηματοπιστωτικές αγορές χαρακτηρίζονται από συνεχή μεταβλητότητα. Η παρακολούθηση των *red flags* πραγματοποιείται σε συνεχή βάση, καθώς νέες πληροφορίες, οικονομικά αποτελέσματα, ειδησεογραφικές εξελίξεις ή αλλαγές στις συνθήκες της αγοράς μπορούν να τροποποιήσουν την εικόνα κινδύνου μιας εταιρείας. Σε αυτό το πλαίσιο, τα σύγχρονα LLM μπορούν να διαδραματίσουν σημαντικό ρόλο, αναλύοντας σε πραγματικό χρόνο οικονομικές ανακοινώσεις, εταιρικές αναφορές, ειδήσεις και συναισθηματική ένταση των αγορών, προσφέροντας στους αναλυτές ένα αυτοματοποιημένο εργαλείο έγκαιρης προειδοποίησης.

Συνολικά, το *red flagging* αποτελεί ένα κείμενο στοιχείο της διαδικασίας διαχείρισης κινδύνων, ενώ η αξιοποίησή του μέσω LLM μπορεί να ενισχύσει σημαντικά την αποτελεσματικότητα της διαχείρισης κινδύνων, οδηγώντας σε πιο ακριβείς, τεκμηριωμένες και έγκαιρες επενδυτικές αποφάσεις.

### (3.7.α Περιγραφή Εφαρμογής Red Flagging)

Για την ανίχνευση γλωσσικών προειδοποιητικών σημάτων σε εταιρικές ανακοινώσεις και άλλα κείμενα αναπτύχθηκε μια εφαρμογή red flagging που λειτουργεί με το τοπικό Llama 3 μέσω Ollama. Ο στόχος της εφαρμογής είναι να εντοπίζει φράσεις που ενδέχεται να είναι ασαφείς, υπερβολικά θετικές ή παραπλανητικές, ώστε να υποστηρίζει αναλυτές στη διαδικασία risk assessment.

Η διεπαφή δημιουργήθηκε με streamlit, ξεκινώντας με τη ρύθμιση των παραμέτρων της σελίδας και την εισαγωγή τίτλου και περιγραφής.

```
st.set_page_config(page_title="Red Flag Detector", page_icon="🚩")
st.title("🚩 Corporate Announcement Red Flag Detector")
st.write("Analyze a company's statement for vague, overly-positive, or misleading language using a local LLaMA 3 model via Ollama.")
```

Ο χρήστης εισάγει το όνομα της εταιρείας, την ημερομηνία της ανακοίνωσης και το πλήρες κείμενο της εταιρικής δήλωσης.

```
company = st.text_input("Company name", "Starbucks Corporation")
date = st.text_input("Announcement date", "2024-04-30")

announcement = st.text_area(
    "Paste the corporate announcement text below:",
    """In a highly challenged environment, this quarter's results do not reflect the power of our brand, our capabilities or the opportunities ahead. We have a clear plan to execute and the entire organization is mobilized around it. We are very confident in our long-term and know that our Triple Shot Reinvention with Two Pumps strategy will deliver on the limitless potential of this brand."""
)
```

Για την κατηγοριοποίηση κάθε red flag χρησιμοποιήθηκε μια συνάρτηση που εξετάζει το περιεχόμενο της αιτιολόγησης και επιστρέφει χρώμα ανάλογα με τον τύπο κινδύνου. Η λογική βασίζεται σε λέξεις-κλειδιά που συνδέονται με ασάφεια (vague), υπερβολική θετικότητα (overly positive) ή παραπλανητικές διατυπώσεις (misleading).

```
def get_highlight_color(reason: str) -> str:
    reason_lower = reason.lower()
    if any(k in reason_lower for k in ["vague", "unclear", "nonspecific", "ambiguous"]):
        return "#fff3cd" # yellow
    elif any(k in reason_lower for k in ["positive", "confident", "optimistic", "hyperbolic", "boastful"]):
        return "#f8d7da" # red/pink
    elif any(k in reason_lower for k in ["misleading", "unrealistic", "exaggerated", "unsupported", "deceptive"]):
        return "#ffe5b4" # orange
    else:
        return "#e2e3e5" # gray (other)
```

Η επισήμανση των προβληματικών φράσεων στο κείμενο υλοποιήθηκε με αντικατάσταση των εμφανίσεών τους από HTML που εφαρμόζει χρωματισμό, έντονη γραφή και αιτιολόγηση.

```
def highlight_phrases(text, flags):
    flags_sorted = sorted(flags, key=lambda x: -len(x["phrase"]))
    for f in flags_sorted:
        phrase = re.escape(f["phrase"])
        color = get_highlight_color(f["reason"])
        tooltip = html.escape(f.get("justification", f["reason"]))
        text = re.sub(
            phrase,
            lambda m: f"<span style='background-color:{color}; color:#000; font-weight:bold;'
title='{tooltip}'>{m.group(0)}</span>",
            text,
            flags=re.IGNORECASE
        )
    return text
```

Στη συνέχεια δημιουργείται το prompt για το Llama 3. Το prompt καθοδηγεί το μοντέλο να λειτουργήσει ως αναλυτής εταιρικών επικοινωνιών και να επιστρέψει αποκλειστικά JSON, χωρίς επιπλέον κείμενο. Το prompt προσαρμόζεται δυναμικά με βάση τα δεδομένα του χρήστη.

```
prompt = f"""
You are a financial communications analyst.
Read the following corporate announcement from {company} dated {date}.
Identify any vague, overly-positive, or potentially misleading phrases.
Return *only* a valid JSON list of objects...
Announcement:
\{announcement}
"""
payload = {"model": "Llama 3", "prompt": prompt, "stream": True}
```

Η αίτηση αποστέλλεται στο API του Ollama με ενεργοποιημένο streaming, ώστε η απάντηση να παραληφθεί τμηματικά. Το κείμενο συλλέγεται σταδιακά μέχρι να γίνει πλήρες.

```
response = requests.post(API_URL, json=payload, stream=True)
collected_text = ""
for line in response.iter_lines(decode_unicode=True):
    if line:
        data = json.loads(line)
        collected_text += data.get("response", "")
```

Το μοντέλο επιστρέφει μια λίστα που περιέχει τις φράσεις, τον τύπο του red flag και την αιτιολόγηση. Το JSON εξάγεται από το κείμενο με χρήση κανονικής έκφρασης.

```
match = re.search(r'\[.*?\]', result_text, re.DOTALL)
flags = json.loads(match.group(0))
```

Το τελικό τμήμα της εφαρμογής αφορά την προβολή και παρουσίαση των αποτελεσμάτων στο περιβάλλον streamlit. Το κείμενο εμφανίζεται με χρωματισμένες επισημάνσεις και ο χρήστης μπορεί να επεκτείνει κάθε flag για να δει λεπτομέρειες, ενώ παρέχεται και επιλογή λήψης του JSON αρχείου.

```
highlighted_html = highlight_phrases(announcement, flags).replace("\n", "<br>")

st.subheader("🔍 Highlighted Announcement")
st.markdown(
    f"<div style='max-height:400px; overflow:auto; line-height:1.5;'>{highlighted_html}</div>",
    unsafe_allow_html=True
)

st.markdown("""
**Color Legend:**
🟡 Vague / unclear     🔴 Overly positive     🟠 Misleading / unrealistic     ⬜ Other
""")

st.subheader("🚩 Detected Red Flags")
for f in flags:
    color = get_highlight_color(f["reason"])
    justification = f.get("justification", "No justification provided.")
    with st.expander(f"⚠️ {f['phrase']} ({f['reason'].capitalize()})"):
        st.markdown(
            f"<div style='background-color:{color}; padding:6px; border-radius:4px;'>
            f"<strong>Reason:</strong> {f['reason'].capitalize()}<br>"
            f"<strong>Justification:</strong> {justification}"
            f"</div>",
            unsafe_allow_html=True
        )

# --- Download JSON ---
st.download_button(
    "📄 Download Red Flags",
    json.dumps(flags, indent=2),
    file_name="red_flags.json",
    mime="application/json"
)
```

### (3.7.β Παράδειγμα Εφαρμογής Red Flagging)

Αναζητούνται red flags σε μια ανακοίνωση της διεύθυνσης των Starbucks:

«Σε ένα περιβάλλον με μεγάλες προκλήσεις, τα αποτελέσματα αυτού του τριμήνου δεν αντικατοπτρίζουν τη δύναμη του εμπορικού μας σήματος, τις ικανότητές μας ή τις ευκαιρίες που έχουμε μπροστά μας. Διαθέτουμε ένα σαφές σχέδιο προς εκτέλεση και ολόκληρος ο οργανισμός είναι κινητοποιημένος γύρω από αυτό. Είμαστε πολύ αισιόδοξοι για το μακροπρόθεσμο μέλλον μας και γνωρίζουμε ότι η στρατηγική μας «Triple Shot Reinvention με Two Pumps» θα αξιοποιήσει το απεριόριστο δυναμικό αυτού του εμπορικού σήματος.»

Για τη συζήτηση των αποτελεσμάτων ανατρέξτε στην παράγραφο 4.5.

**Corporate Announcement Red Flag Detector**

Analyze a company's statement for vague, overly-positive, or misleading language using a local LLaMA 3 model via Ollama.

Company name  
Starbucks Corporation

Announcement date  
2024-04-30

Paste the corporate announcement text below:  
In a highly challenged environment, this quarter's results do not reflect the power of our brand, our capabilities or the opportunities ahead. We have a clear plan to execute and the entire organization is mobilized around it We are very confident in our long-term and know that our Triple Shot

Analyze

✓ Analysis complete!

🔍 Highlighted Announcement  
In a highly challenged environment, this quarter's results do not reflect the power of our brand, our capabilities or the opportunities ahead. We have a clear plan to execute and the entire organization is mobilized around it. We are very confident in our long-term and know that our Triple Shot Reinvention with Two Pumps strategy will deliver on the limitless potential of this brand.

Color Legend:  
● Vague / unclear ● Overly positive ● Misleading / unrealistic ● Other

**Detected Red Flags**  
⚠ this quarter's results do not reflect the power of our brand, (Vague)

⚠ our capabilities or the opportunities ahead. (Vague)

⚠ We have a clear plan to execute and the entire organization is mobilized around it. (Overly positive)

⚠ We are very confident in our long-term (Overly positive)

⚠ know that our Triple Shot Reinvention with Two Pumps strategy will deliver on the limitless potential of this brand. (Misleading)

 Download Red Flags

#### Downloaded json with Red Flags

```
[
  {
    "phrase": "this quarter results do not reflect the power of our brand,",
    "reason": "vague",
    "justification": "The phrase 'do not reflect' is overly general and doesn't provide specific details about what went wrong or how it will be improved."
  },
  {
    "phrase": "our capabilities or the opportunities ahead.",
    "reason": "vague",
    "justification": "This sentence lacks concrete specifics about what these capabilities and opportunities are, making it difficult to understand their impact on the company's performance."
  },
  {
    "phrase": "We have a clear plan to execute and the entire organization is mobilized around it.",
    "reason": "overly positive",
    "justification": "The use of 'clear' and 'mobilized' creates an overly optimistic tone, which may not accurately reflect the complexity or challenges involved in executing the plan."
  },
  {
    "phrase": "We are very confident in our long-term",
    "reason": "overly positive",
    "justification": "The use of 'very confident' is an exaggeration and may raise skepticism about the company's ability to deliver on its claims."
  },
  {
    "phrase": "know that our Triple Shot Reinvention with Two Pumps strategy will deliver on the limitless potential of this brand.",
    "reason": "misleading",
    "justification": "The phrase 'limitless potential' is an unrealistic and unquantifiable claim, which may not accurately reflect the company's actual performance or prospects."
  }
]
```

```
]
```

## Μετάφραση json με Red Flags

```
[  
  {  
    "φράση": "τα αποτελέσματα αυτού του τριμήνου δεν αντικατοπτρίζουν τη δύναμη της μάρκας μας," ,  
    "λόγος": "αόριστο",  
    "δικαιολόγηση": "Η φράση 'δεν αντικατοπτρίζουν' είναι υπερβολικά γενική και δεν παρέχει συγκεκριμένες λεπτομέρειες για το τι πήγε λάθος ή πώς θα βελτιωθεί."  
  },  
  {  
    "φράση": "τις δυνατότητές μας ή τις ευκαιρίες που έχουμε μπροστά μας.",  
    "λόγος": "αόριστο",  
    "δικαιολόγηση": "Η πρόταση αυτή δεν παρέχει συγκεκριμένες λεπτομέρειες για το ποιες είναι αυτές οι δυνατότητες και ευκαιρίες, καθιστώντας δύσκολη την κατανόηση της επίδρασής τους στην απόδοση της εταιρείας."  
  },  
  {  
    "φράση": "Έχουμε ένα σαφές σχέδιο για υλοποίηση και ολόκληρη η οργάνωση έχει κινητοποιηθεί γύρω από αυτό.",  
    "λόγος": "υπερβολικά θετικό",  
    "δικαιολόγηση": "Η χρήση των λέξεων 'σαφές' και 'κινητοποιηθεί' δημιουργεί έναν υπερβολικά αισιόδοξο τόνο, ο οποίος μπορεί να μην αντικατοπτρίζει με ακρίβεια την πολυπλοκότητα ή τις προκλήσεις που εμπλέκονται στην υλοποίηση του σχεδίου."  
  },  
  {  
    "φράση": "Είμαστε πολύ σίγουροι για τη μακροπρόθεσμη",  
    "λόγος": "υπερβολικά θετικό",  
    "δικαιολόγηση": "Η χρήση του 'πολύ σίγουροι' αποτελεί υπερβολή και μπορεί να προκαλέσει σκεπτικισμό σχετικά με την ικανότητα της εταιρείας να εκπληρώσει τις δεσμεύσεις της."  
  },  
  {  
    "φράση": "γνωρίζουμε ότι η στρατηγική μας Triple Shot Reinvention with Two Pumps θα αξιοποιήσει το απεριόριστο δυναμικό αυτής της μάρκας.",  
    "λόγος": "παραπλανητικό",  
    "δικαιολόγηση": "Η φράση 'απεριόριστο δυναμικό' είναι μια μη ρεαλιστική και μη μετρήσιμη δήλωση, που μπορεί να μην αντικατοπτρίζει με ακρίβεια την πραγματική απόδοση ή προοπτικές της εταιρείας."  
  }  
]
```

## ΚΕΦΑΛΑΙΟ 4 Συζήτηση Αποτελεσμάτων

### ( 4.1 Συζήτηση Εργαλείων Εφαρμογών)

Στην παρούσα εργασία αξιοποιήθηκαν διαφορετικά εργαλεία για την ανάπτυξη εφαρμογών βασισμένων σε μεγάλα γλωσσικά μοντέλα (LLM) στον τομέα του χρηματιστηρίου. Κάθε εργαλείο συνέβαλε με διαφορετικό τρόπο στη συλλογή, ανάλυση και οπτικοποίηση δεδομένων, προσφέροντας ένα ολοκληρωμένο περιβάλλον διερεύνησης της αγοράς.

Το streamlit αποτέλεσε τον βασικό μηχανισμό υλοποίησης του γραφικού περιβάλλοντος και διαδραστικής παρουσίασης των δεδομένων. Διευκόλυε σημαντικά την ενσωμάτωση τόσο ποσοτικών δεδομένων (ιστορικές τιμές, τεχνικοί δείκτες, όγκοι συναλλαγών) όσο και ποιοτικών πληροφοριών (ειδήσεις, αναλύσεις LLM). Μέσω των διαθέσιμων widgets, όπως επιλογείς μετοχών, sliders χρονικών διαστημάτων και διαδραστικά γραφήματα, ο χρήστης μπορούσε εύκολα να προσαρμόσει την ανάλυση και να παράγει δικές του οπτικοποιήσεις. Η δυνατότητα απλής και γρήγορης ανάπτυξης χωρίς ανάγκη σύνθετης εγκατάστασης το καθιστά ιδιαίτερα κατάλληλο για πειραματικές εφαρμογές. Ωστόσο, η πλατφόρμα εμφανίζει περιορισμούς σε μεγαλύτερης κλίμακας έργα ή περιπτώσεις όπου απαιτείται υψηλή απόδοση και ταυτόχρονη χρήση από πολλούς χρήστες.

Η ενσωμάτωση του finvizfinance προσέφερε ένα αξιόπιστο σύστημα άντλησης ειδησεογραφίας για συγκεκριμένες μετοχές. Τα δεδομένα ειδήσεων προβάλλονταν μέσω του streamlit, ενώ στη συνέχεια επεξεργάζονταν από το LLM Llama 3 για την εξαγωγή ανάλυσης συναισθήματος. Με αυτόν τον τρόπο, ο χρήστης μπορούσε να εξαγει γρήγορα συμπεράσματα σχετικά με τον αντίκτυπο των ειδήσεων, χωρίς να απαιτείται η μελέτη πλήρων άρθρων. Παρόλα αυτά, ο περιορισμός των έως 100 διαθέσιμων ειδήσεων ανά κλήση αποτελεί σημαντικό μειονέκτημα, ειδικά για μετοχές με υψηλή ειδησεογραφική δραστηριότητα ή για εφαρμογές πρόβλεψης που απαιτούν μεγάλο όγκο ιστορικού υλικού.

Αντίθετα, το yfinance προσέφερε πλήρη και ανεξάντλητη πρόσβαση σε ιστορικά δεδομένα τιμών, καθιστώντας το ιδανικό για αναλύσεις μεγάλης χρονικής διάρκειας. Η ευκολία στην άντληση δεδομένων και η άμεση ενσωμάτωσή τους στην εφαρμογή επέτρεψαν την παραγωγή χρονοσειρών, τεχνικών δεικτών, συγκριτικών γραφημάτων και άλλων στοιχείων απαραίτητων για ποσοτική ανάλυση. Η απουσία περιορισμών στον όγκο δεδομένων καθιστά το yfinance σαφώς πιο αξιόπιστο για στρατηγικές πρόβλεψης και ανάλυσης σε σχέση με πηγές που επιβάλλουν όρια, όπως το finviz.

Το Llama 3 αποτέλεσε κεντρικό στοιχείο της εφαρμογής, καθώς χρησιμοποιήθηκε για την επεξεργασία, ανάλυση και ερμηνεία τόσο ποσοτικών όσο και ποιοτικών δεδομένων. Η απόδοση του Llama3 θα αναλυθεί εκτενώς και στις επόμενες παραγράφους για κάθε μία από τις εφαρμογές που αναπτύχθηκαν.

Συνολικά, ο συνδυασμός streamlit, finvizfinance, yfinance και Llama 3 δημιούργησε ένα πολυδιάστατο και πλήρως ενοποιημένο σύστημα χρηματιστηριακής ανάλυσης, το οποίο αξιοποιεί τη συμπληρωματικότητα των επιμέρους εργαλείων για να καλύψει όλο το φάσμα της πληροφορίας και της υποστήριξης λήψης αποφάσεων.

## (4.2 Συζήτηση Αποτελεσμάτων Πρόβλεψης Απόδοσης Μετοχής)

---

Στην εφαρμογή πρόβλεψης απόδοσης μετοχής της παραγράφου 3.4 αξιοποιήθηκαν πέντε διαφορετικοί αλγόριθμοι μηχανικής μάθησης: Random Forest, Extra Trees, Support Vector Regression (SVR), Multilayer Perceptron (MLP) και Gradient Boosting.

Η αξιολόγησή τους βασίζεται σε τέσσερις δείκτες αξιολόγησης (metrics): Mean Absolute Error (MAE), Mean Squared Error (MSE), συντελεστή προσδιορισμού  $R^2$  και κατευθυντική ακρίβεια (Directional Accuracy). Οι συγκεκριμένοι δείκτες επιλέχθηκαν ώστε να αποτυπώσουν διαφορετικές πτυχές της απόδοσης των μοντέλων, από το μέσο μέγεθος του σφάλματος έως την ικανότητα πρόβλεψης της τάσης.

### Μέσο Απόλυτο Σφάλμα (Mean Absolute Error, MAE)

---

Το MAE μετράει τη μέση απόκλιση μεταξύ των προβλεπόμενων και των πραγματικών τιμών.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$y$  = πραγματική τιμή

$\hat{y}$  = προβλεπόμενη τιμή από το μοντέλο

$n$  = πλήθος παρατηρήσεων

Όσο μικρότερο το MAE, τόσο πιο ακριβές το μοντέλο.

### Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error, MSE)

---

Το MSE υπολογίζει τη μέση τετραγωνική απόκλιση μεταξύ των προβλεπόμενων και των πραγματικών τιμών, είναι δηλαδή πιο ευαίσθητο σε μεγάλα σφάλματα λόγω του τετραγωνικού όρου.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Όσο μεγαλύτερο είναι το MSE, τόσο χειρότερη η απόδοση του μοντέλου. Η τετραγωνική φύση του σφάλματος σημαίνει ότι τα μεγάλα λάθη επηρεάζουν σημαντικά την τιμή, καθιστώντας το MSE ευαίσθητο σε outliers.

### Συντελεστής $R^2$

---

Ο  $R^2$  αποτυπώνει το ποσοστό της διακύμανσης της μεταβλητής-στόχου που εξηγείται από το μοντέλο.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$\bar{y}$  = η μέση τιμή

Τιμές κοντά στο 1 υποδηλώνουν ισχυρή προσαρμογή, τιμές κοντά στο 0 ισοδυναμούν με βασική απόδοση, ενώ αρνητικές τιμές δείχνουν χειρότερη επίδοση από την απλή πρόβλεψη του μέσου όρου. Αρνητικές τιμές σημαίνουν ότι το μοντέλο είναι χειρότερο από τη βασική πρόβλεψη με μέση τιμή.

## Κατευθυντική Ακρίβεια (Directional Accuracy)

Μετρά το ποσοστό των περιπτώσεων όπου το μοντέλο προβλέπει σωστά την κατεύθυνση της μεταβολής (π.χ. αν θα αυξηθεί ή θα μειωθεί η τιμή).

Ο Πίνακας 4.1 παρουσιάζει τα αποτελέσματα αξιολόγησης των πέντε αλγορίθμων πρόβλεψης μετοχών, μετρώντας τόσο την απόλυτη και τετραγωνική απόκλιση των προβλέψεων (MAE, MSE) όσο και την ικανότητά τους να προσαρμόζονται στα δεδομένα ( $R^2$ ) και να εντοπίζουν σωστά την κατεύθυνση της μεταβολής της τιμής (Directional Accuracy).

Πίνακας 4.1: Αποτελέσματα αξιολόγησης αλγορίθμων πρόβλεψης μετοχών

	Mean Absolute Error (MAE)	Mean Squared Error (MSE)	$R^2$ Score	Directional Accuracy (%)
Random Forest	0.83	1.04	0.02	58.18
Extra Trees	0.88	1.27	-0.19	52.73
SVR	0.79	1.03	0.03	54.55
MLP	0.99	1.43	-0.34	58.18
Gradient Boosting	0.82	0.99	0.07	61.82

Τα πέντε μοντέλα εμφάνισαν διαφοροποιημένες επιδόσεις, τόσο ως προς το μέγεθος των σφαλμάτων όσο και ως προς την ικανότητά τους να προβλέπουν την κατεύθυνση των μεταβολών.

Το Gradient Boosting αναδείχθηκε το πιο αποδοτικό και ισορροπημένο μοντέλο. Παρουσίασε τη χαμηλότερη τιμή MSE (0.99), τη μεγαλύτερη τιμή  $R^2$  (0.07) και την υψηλότερη κατευθυντική ακρίβεια (61.82%). Αυτό υποδηλώνει ότι καταφέρνει να συλλάβει αποτελεσματικά τις σχέσεις μεταξύ των μεταβλητών και να προβλέψει σωστά την τάση της μετοχής.

Το SVR πέτυχε τη μικρότερη τιμή MAE (0.79), γεγονός που δείχνει ότι οι προβλέψεις του βρίσκονται γενικά πιο κοντά στις πραγματικές τιμές, αν και υστερεί σε σχέση με το Gradient Boosting στην πρόβλεψη κατεύθυνσης.

Το Random Forest είχε σταθερή συμπεριφορά με μέτρια σφάλματα και θετικό  $R^2$ , ενώ τα Extra Trees και MLP εμφάνισαν αρνητικές τιμές  $R^2$ , κάτι που υποδηλώνει ότι δεν ξεπέρασαν την απόδοση ενός απλού βασικού μοντέλου. Παρ' όλα αυτά, το MLP διατήρησε κατευθυντική ακρίβεια αντίστοιχη με εκείνη του Random Forest, στοιχείο που δείχνει ότι, παρά την αδύναμη παλινδρόμηση, εντοπίζει μερικώς την τάση.

Οι χαμηλές τιμές  $R^2$  που παρατηρούνται συνολικά (από -0.34 έως 0.07) δεν αποτελούν ένδειξη αποτυχίας των μοντέλων, αλλά αντανάκλαση της πολυπλοκότητας και του υψηλού θορύβου που χαρακτηρίζει τις χρηματοπιστωτικές αγορές. Σε τέτοιου είδους δεδομένα, στόχος δεν είναι η ακριβής πρόβλεψη του μεγέθους των μεταβολών, αλλά η στατιστικά καλύτερη από την τυχαία πρόβλεψη της κατεύθυνσης. Τα αποτελέσματα των μοντέλων μπορούν να συγκριθούν με αντίστοιχες μελέτες στη βιβλιογραφία [13], όπου παρατηρείται χαμηλότερη απόδοση, γεγονός που υπογραμμίζει την σχετική αποτελεσματικότητα των αλγορίθμων που χρησιμοποιήθηκαν στην παρούσα εργασία.

Η κατευθυντική ακρίβεια πάνω από 55% θεωρείται στατιστικά σημαντική σε δεδομένα χρηματιστηριακών αγορών, τα οποία χαρακτηρίζονται από υψηλό θόρυβο και τυχαίες διακυμάνσεις. Ακόμη και μικρή υπεροχή έναντι της τυχαίας πρόβλεψης (50%) μεταφράζεται σε σημαντικό πλεονέκτημα όταν χρησιμοποιείται σε συνδυασμό με κατάλληλες στρατηγικές διαχείρισης κινδύνου.

## ( 4.3 Συζήτηση Αποτελεσμάτων Σύνοψης & Ανάλυσης Αναφορών)

Από την ανάπτυξη και τα αποτελέσματα της εφαρμογής αναδείχθηκαν τα πλεονεκτήματα και οι προκλήσεις της χρήσης του Llama 3 για σύνοψη και ανάλυση οικονομικών αναφορών.

### Επάρκεια και Ακρίβεια αποτελεσμάτων

Τόσο η γρήγορη όσο και η αναλυτική εκδοχή των περιλήψεων είναι γενικά ακριβείς και επαρκείς, εστιάζοντας στα πιο σημαντικά σημεία της αναφοράς.

Ως προς τα βασικά οικονομικά μεγέθη, ο έλεγχος των απαντήσεων παρουσιάζεται στον Πίνακα 4.2.

Πίνακας 4.2: Έλεγχος απαντήσεων του παραδείγματος στην παρ. 3.4

Ερωτήσεις στα Οικονομικά Μεγέθη		
	Γρήγορη Απάντηση Llama 3	Αναλυτική Απάντηση Llama 3
Πωλήσεις 9Μ24 και σύγκριση με το 2023	Ανεπαρκής	Επαρκής
ΕΒΙΤΔΑ 9Μ24 και σύγκριση με το 2023	Ανεπαρκής	Επαρκής
Καθαρά Κέρδη 9Μ24 και σύγκριση με το 2023	Επαρκής	Επαρκής
Απασχολούμενα κεφάλαια 9Μ24 και σύγκριση με το 2023	Ανεπαρκής	Επαρκής
Καθαρός Δανεισμός 9Μ24 και σύγκριση με το 2023	Επαρκής	Επαρκής
Προσωρινό μέρισμα ανά μετοχή στους μετόχους	Επαρκής	Επαρκής
Επιπτώσεις των τιμών του πετρελαίου	Επαρκής	Επαρκής

Όπως φαίνεται, η αναλυτική εκδοχή παρέχει μεγαλύτερη επάρκεια σε σχέση με τη γρήγορη, ειδικά για στοιχεία που απαιτούν υπολογιστική εξαγωγή ή σύνθετες συγκρίσεις.

### Χρόνος επεξεργασίας (latency)

Η επεξεργασία των αναφορών απαιτεί σημαντικό χρόνο, ιδιαίτερα σε υπολογιστές με περιορισμένες υπολογιστικές δυνατότητες, όπως ένα laptop Intel Core i7 που χρησιμοποιήθηκε.

Πίνακας 4.3: Υπολογιστικός χρόνος για κείμενο 14.180 χαρακτήρων

	Γρήγορη Απάντηση Llama 3 Χρόνος (sec)	Αναλυτική Απάντηση Llama 3 Χρόνος (sec)
Περίληψη Αναφοράς	3616	5829
Απάντηση σε ερωτήσεις για Βασικά Οικονομικά Μεγέθη και Στρατηγική	6020	6035

Παρατηρείται σημαντική αύξηση χρόνου για την αναλυτική εκδοχή. Η αναλυτική επεξεργασία βελτιώνει την ποιότητα των απαντήσεων σε σύνθετα ερωτήματα, ενώ η γρήγορη εκδοχή είναι επαρκής για πιο γενικές ανάγκες.

Η τρέχουσα υλοποίηση χρησιμοποιεί τοπικά τη CPU, γεγονός που περιορίζει τη δυνατότητα παράλληλης επεξεργασίας μεγάλου όγκου chunks. Η χρήση μικρότερων τμημάτων κειμένου (chunks) μειώνει τον χρόνο ανά αίτημα, με μικρή απώλεια πληρότητας, ενώ στο τέλος το μοντέλο συνοψίζει τις επιμέρους περιλήψεις για ολοκληρωμένη εκτίμηση.

### Αριθμητική ακρίβεια

---

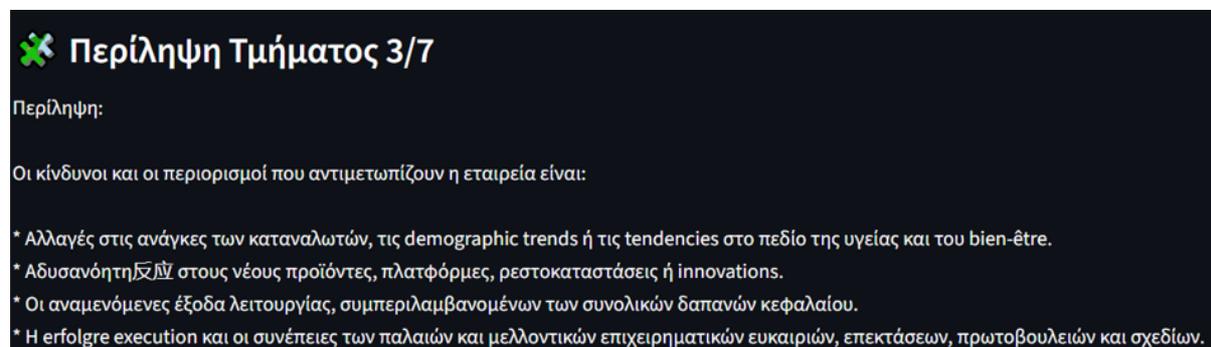
Η αριθμητική ακρίβεια των απαντήσεων είναι γενικά ικανοποιητική για βασικά οικονομικά μεγέθη, όπως εμφανίζεται στον Πίνακα 4.2. Ωστόσο, σε σύνθετα κείμενα με πολλαπλές μονάδες μέτρησης ή διαφορετικά νομίσματα απαιτείται κανονικοποίηση και ανθρώπινη εποπτεία, καθώς τα LLM δεν εγγυώνται απόλυτη ακρίβεια σε αριθμητικά δεδομένα.

### Μεταφραστικές δυνατότητες

---

Η ανάλυση αναφορών σε διαφορετικές γλώσσες αναδεικνύει περιορισμούς στη σωστή μετάφραση οικονομικών όρων. Παρά την ικανότητα επεξεργασίας πολυγλωσσικών κειμένων, η ακριβής μετάφραση ειδικής ορολογίας παραμένει πρόκληση.

Για παράδειγμα, η μετάφραση οικονομικής αναφοράς της εταιρείας Starbucks από τα αγγλικά στα ελληνικά παρουσίασε δυσκολίες σε ορισμένους όρους και μονάδες.



**✖ Περιλήψη Τμήματος 3/7**

Περίληψη:

Οι κίνδυνοι και οι περιορισμοί που αντιμετωπίζουν η εταιρεία είναι:

- \* Αλλαγές στις ανάγκες των καταναλωτών, τις demographic trends ή τις tendencies στο πεδίο της υγείας και του bien-être.
- \* Αδυσανόητη ~~αν~~ στους νέους προϊόντες, πλατφόρμες, ρεστοκαταστάσεις ή innovations.
- \* Οι αναμενόμενες έξοδα λειτουργίας, συμπεριλαμβανομένων των συνολικών δαπανών κεφαλαίου.
- \* Η erfolgreiche execution και οι συνέπειες των παλαιών και μελλοντικών επιχειρηματικών ευκαιριών, επεκτάσεων, πρωτοβουλιών και σχεδίων.

Η εφαρμογή της σύνοψης και ανάλυσης αναφορών με Llama 3 επιβεβαιώνει την αποτελεσματικότητα των LLM στη μείωση χρόνου επεξεργασίας και στη βελτίωση της κατανόησης μεγάλων κειμένων. Ωστόσο, οι περιορισμοί σε υπολογιστικούς πόρους, αριθμητική ακρίβεια και μεταφραστική δυνατότητα καθιστούν αναγκαία την ανθρώπινη εποπτεία, ειδικά για κρίσιμες χρηματοοικονομικές αποφάσεις.

## (4.4 Συζήτηση Αποτελεσμάτων Chatbot)

---

Η ανάπτυξη του chatbot για χρηματιστηριακή ανάλυση αξιοποίησε συστηματικά διάφορες τεχνικές prompt engineering, με στόχο τη βελτίωση της ακρίβειας, της συνάφειας και της επεξηγηματικής ικανότητας του LLM. Στις περισσότερες περιπτώσεις, το σύστημα παράγαγε απαντήσεις υψηλής πληρότητας, ενώ ο χρόνος εκτέλεσης παρέμεινε σταθερός, αν και υψηλός (μερικά λεπτά), λόγω της διαδικτυακής αναζήτησης που προηγείται κάθε απάντησης.

### *Zero Shot Prompting*

---

Το Zero-Shot Prompting, όπου το μοντέλο δέχεται μόνο μία ερώτηση χωρίς πρόσθετη καθοδήγηση ή παραδείγματα, παρείχε ικανοποιητικά αποτελέσματα για απλά ερωτήματα όπως η βασική ταυτοποίηση εταιρειών και η γενική περιγραφή της απόδοσης των μετοχών.

Στην περίπτωση της Apple (AAPL), το μοντέλο ήταν σε θέση να προσδιορίσει σωστά το όνομα της εταιρείας και τον τομέα της, ενώ έδωσε και μια συνοπτική εικόνα της πρόσφατης απόδοσης της μετοχής. Ωστόσο, η ανάλυση ήταν περιορισμένη και λιγότερο επενδυτικά χρήσιμη, καθώς απουσίαζε σε βάθος αξιολόγηση ή ειδικές χρηματοοικονομικές πληροφορίες.

Το Zero-Shot Prompting αναδεικνύει τη σημασία της ακρίβειας στη διατύπωση του prompt: όσο πιο ασαφής είναι η εντολή, τόσο πιο γενική είναι η απάντηση. Παρά τους περιορισμούς, η τεχνική αυτή είναι χρήσιμη όταν απαιτείται ταχύτητα, ευελιξία ή όταν δεν υπάρχουν διαθέσιμα παραδείγματα για το μοντέλο.

### *Few-Shot Prompting*

---

Στο Few-Shot Prompting, όπου παρέχονται στο μοντέλο συγκεκριμένα παραδείγματα (AAPL, TSLA, NVDA), παρατηρήθηκε αισθητή βελτίωση τόσο στην ακρίβεια όσο και στη συνέπεια της απάντησης.

Η ανάλυση της Microsoft (MSFT) ακολούθησε πιστά τη δομή των παραδειγμάτων, με σαφή παρουσίαση του ονόματος, του τομέα και της πρόσφατης χρηματιστηριακής εικόνας. Η απάντηση ήταν περισσότερο οργανωμένη, λεπτομερής και χρήσιμη για επενδυτές.

Το Few-Shot Prompting αποδείχθηκε ιδιαίτερα αποτελεσματικό για εφαρμογές που απαιτούν συγκριτική, δομημένη και συνεπή πληροφόρηση για πολλαπλές μετοχές. Η ποιότητα των παραδειγμάτων που παρέχονται στο μοντέλο αποτελεί κρίσιμο παράγοντα για την ποιότητα της παραγόμενης ανάλυσης.

### *Chain of Thought*

---

Η τεχνική του Chain of Thought Prompting επιτρέπει στο μοντέλο γλώσσας να σκέφτεται βήμα-βήμα, παρέχοντας μια πιο αναλυτική και τεκμηριωμένη απάντηση σε σύνθετα ερωτήματα. Στην περίπτωση της NVIDIA, η διαδικασία οδήγησε το μοντέλο να εξετάσει πρώτα τον κύριο τομέα δραστηριότητας και τις βασικές επιχειρηματικές μονάδες της εταιρείας, στη συνέχεια να αναλύσει τις τάσεις της αγοράς και του κλάδου, να αξιολογήσει πρόσφατες εταιρικές εξελίξεις και να λάβει υπόψη πιθανούς κινδύνους και μεταβλητότητα. Η δομημένη αυτή προσέγγιση επιτρέπει την παραγωγή μιας ολοκληρωμένης και ισορροπημένης επενδυτικής πρότασης.

Το πλεονέκτημα του Chain of Thought Prompting είναι ότι μειώνει την πιθανότητα παραλείψεων ή γενικών απαντήσεων, καθώς το μοντέλο αναγκάζεται να αιτιολογήσει κάθε βήμα της σκέψης του

πριν καταλήξει σε συμπέρασμα. Παράλληλα, επιτρέπει στους χρήστες να κατανοήσουν τη λογική πίσω από την πρόταση επένδυσης, γεγονός που αυξάνει τη διαφάνεια και την αξιοπιστία της ανάλυσης. Στην περίπτωση της NVIDIA, αυτή η προσέγγιση ανέδειξε τόσο τις ευκαιρίες που προσφέρει η αυξανόμενη ζήτηση για προϊόντα AI όσο και τους κινδύνους που σχετίζονται με την επιβράδυνση των κερδών και τη μεταβλητότητα του κλάδου.

Το Chain of Thought Prompting αποτελεί ένα ισχυρό εργαλείο για την παραγωγή πολυδιάστατων και εμπειριστατωμένων αναλύσεων. Η μέθοδος αυτή αναδεικνύει τη σημασία της καθοδήγησης του μοντέλου μέσω βήμα-βήμα σκέψης, αυξάνοντας την ακρίβεια, τη συνέπεια και την αξία των παραγόμενων αποτελεσμάτων, ιδιαίτερα σε σύνθετες χρηματοοικονομικές αναλύσεις.

### *Self-Consistency*

---

Η ανάλυση των τριών ανεξάρτητων διαδρομών συλλογισμού για την επένδυση στην Tesla υπογραμμίζει τη σημασία της τεχνικής του prompt engineering στη λήψη ποιοτικών και χρήσιμων αποτελεσμάτων. Το αρχικό prompt καθορίζει με σαφήνεια τις προσδοκίες, ζητώντας από το μοντέλο να εξετάσει τα οικονομικά στοιχεία, τις πρόσφατες τάσεις της αγοράς και τους πιθανούς κινδύνους, καταλήγοντας σε μια σύντομη επενδυτική σύσταση. Η σαφήνεια αυτή εξασφαλίζει ότι οι απαντήσεις είναι συγκροτημένες και επί του θέματος, αποφεύγοντας γενικεύσεις ή εκτροπές.

Η απαίτηση για τρία ανεξάρτητα μονοπάτια σκέψης δείχνει πώς η σωστή διατύπωση του prompt μπορεί να οδηγήσει σε πολλαπλές προοπτικές για το ίδιο ζήτημα, επιτρέποντας τη σύγκριση διαφορετικών λογικών και την αξιολόγηση της αξιοπιστίας τους. Η δομή καθοδήγησης των βημάτων από θεμελιώδη σε τάσεις της αγοράς, στη συνέχεια σε κινδύνους και τέλος σε συστάσεις διασφαλίζει ότι η ανάλυση είναι πλήρης και λογικά οργανωμένη.

Τα διαφορετικά μονοπάτια αναδεικνύουν πτυχές όπως η ανάλυση των θεμελιωδών στοιχείων, η επίδραση της μεταβλητότητας της αγοράς και η σημασία του ρυθμιστικού περιβάλλοντος, δείχνοντας ότι ένα καλά σχεδιασμένο prompt μπορεί να διερευνήσει πολλαπλές όψεις ενός σύνθετου θέματος. Χωρίς την καθοδήγηση αυτή, οι απαντήσεις του μοντέλου θα μπορούσαν να είναι μονοδιάστατες ή ασαφείς.

Η μελέτη των reasoning paths καταδεικνύει ότι το prompt engineering αποτελεί κρίσιμο εργαλείο για την παραγωγή αξιόπιστων και πολυδιάστατων αναλύσεων. Η σαφήνεια, η δομή και η χρήση πολλαπλών παραμέτρων ανάλυσης αυξάνουν σημαντικά την ποιότητα των αποτελεσμάτων και καθιστούν το μοντέλο πιο χρήσιμο για την υποστήριξη επενδυτικών αποφάσεων.

### *Re-Act*

---

Η τεχνική ReAct συνδυάζει δύο βασικές ικανότητες των μοντέλων γλώσσας: τη λογική σκέψη και την εκτέλεση ενεργειών, όπως η αναζήτηση ή η συλλογή δεδομένων. Στο παράδειγμα με την Apple, η διαδικασία ξεκινά με τη σκέψη (Reason), όπου το μοντέλο σκέφτεται πώς να αναλύσει την πορεία της μετοχής, δηλαδή ότι χρειάζεται ιστορικά δεδομένα και ενημερωμένα στοιχεία τιμής. Στη συνέχεια ακολουθεί η ενέργεια (Action), όπου το μοντέλο ορίζει την ενέργεια που πρέπει να κάνει, δηλαδή την άντληση δεδομένων από πηγές όπως το Yahoo Finance.

Με το βήμα παρατήρησης (Observation), το μοντέλο καταγράφει τα αποτελέσματα της ενέργειας, δηλαδή την τρέχουσα τιμή κλεισίματος και τα ιστορικά υψηλά και χαμηλά της μετοχής. Στο Reasoning, αναλύει αυτά τα δεδομένα για να εκτιμήσει τις τάσεις, ενώ στην τελική απάντηση (Answer) συνοψίζει την αξιολόγηση για τον επενδυτή. Η δομή αυτή επιτρέπει να προκύψει μια

απάντηση που είναι τόσο εμπειριστωμένη όσο και τεκμηριωμένη, καθώς βασίζεται σε πραγματικά δεδομένα και λογική ανάλυση.

Η προσέγγιση ReAct είναι ιδιαίτερα χρήσιμη όταν η απάντηση απαιτεί αλληλεπίδραση με εξωτερικές πηγές ή εκτέλεση συγκεκριμένων ενεργειών. Συνδυάζει τη σκέψη με την πράξη, μειώνοντας την πιθανότητα λανθασμένων ή γενικών απαντήσεων και προσφέροντας πιο πρακτικά και αξιοποιήσιμα αποτελέσματα για τον χρήστη. Στο παράδειγμα της AAPL, η τεχνική επιτρέπει την ταχεία αξιολόγηση της τάσης της μετοχής και παρέχει σαφή καθοδήγηση για περαιτέρω ανάλυση ή επενδυτική στρατηγική.

Το ReAct ενισχύει την ακρίβεια και την αποτελεσματικότητα των απαντήσεων, συνδυάζοντας σκέψη και δράση, και αποτελεί ένα ισχυρό εργαλείο για σύνθετες χρηματοοικονομικές αναλύσεις που απαιτούν πρόσβαση σε δεδομένα σε πραγματικό χρόνο.

### *Prompt Chaining*

---

Η τεχνική του Prompt Chaining αποδεικνύει πόσο σημαντικό είναι να σπάει κανείς μια σύνθετη εργασία σε διαδοχικά, λογικά βήματα για να μεγιστοποιηθεί η ακρίβεια και η χρησιμότητα των αποτελεσμάτων ενός μοντέλου γλώσσας. Στο παράδειγμα της Tesla, η διαδικασία χωρίστηκε σε τρία βήματα: σύνοψη των οικονομικών αποτελεσμάτων, ανάλυση των τάσεων της αγοράς και τελική επενδυτική σύσταση. Κάθε βήμα είχε σαφώς καθορισμένο στόχο και εισροές, γεγονός που επέτρεψε στο μοντέλο να εστιάσει σε συγκεκριμένα δεδομένα και να παράγει συγκροτημένη, πλήρη και λογικά δομημένη ανάλυση.

Η χρήση του Prompt Chaining προσφέρει πολλαπλά πλεονεκτήματα. Μειώνει τον κίνδυνο παραλείψεων ή ασάφειας, καθώς κάθε βήμα καθοδηγεί το μοντέλο να επικεντρωθεί σε συγκεκριμένες πληροφορίες. Επιτρέπει την ενσωμάτωση διαφορετικών πτυχών ενός θέματος, όπως τα οικονομικά μεγέθη, οι τάσεις της αγοράς και οι επενδυτικοί κίνδυνοι, με τρόπο που είναι πιο διαφανής και αξιοποιήσιμος από τον χρήστη. Διευκολύνει τη διαδικασία ελέγχου και επικύρωσης των αποτελεσμάτων, καθώς κάθε στάδιο μπορεί να αξιολογηθεί ανεξάρτητα πριν προχωρήσει στο επόμενο.

Στην περίπτωση της Tesla, το chaining επέτρεψε τη δημιουργία μιας ολοκληρωμένης επενδυτικής ανάλυσης που συνδυάζει στοιχεία από οικονομικά δεδομένα, τάσεις αγοράς και στρατηγική αξιολόγηση, παρέχοντας έτσι στον επενδυτή μια σαφή εικόνα τόσο των ευκαιριών όσο και των κινδύνων. Αυτό δείχνει ότι το Prompt Chaining είναι ιδιαίτερα χρήσιμο για σύνθετες εργασίες που απαιτούν πολλαπλά βήματα σκέψης ή ανάλυσης, καθώς βελτιώνει την ακρίβεια, την πληρότητα και την επενδυτική αξία των αποτελεσμάτων.

### *Automatic Prompt Optimization (APO)*

---

Το παράδειγμα Automatic Prompt Optimization (APO) δείχνει πώς η διατύπωση ενός prompt μπορεί να επηρεάσει άμεσα την ποιότητα, τη σαφήνεια και τη χρησιμότητα των παραγόμενων απαντήσεων. Στην περίπτωση της Tesla, δημιουργήθηκαν τρεις εναλλακτικές προσεγγίσεις: η αναδιατύπωση (Reframe), που επαναδιατυπώνει το αρχικό prompt με σαφή και άμεσο τρόπο, η επέκταση (Expand), που ζητά μεγαλύτερη ανάλυση και περισσότερες λεπτομέρειες και ο περιορισμός (Narrow), που περιορίζει την εστίαση σε συγκεκριμένα στοιχεία όπως έσοδα, περιθώρια κέρδους και ταμειακή ροή.

Η εκδοχή αναδιατύπωσης (Reframe) αποδεικνύεται πιο αποτελεσματική, καθώς διατηρεί την επενδυτική συνάφεια και τη συνοπτικότητα περιλαμβάνοντας όμως κρίσιμα σημεία για τον

επενδυτή. Η επέκταση (Expand) προσφέρει περισσότερες λεπτομέρειες, αλλά μπορεί να είναι υπερφορτωμένη ή να περιλαμβάνει στοιχεία που δεν είναι κρίσιμα για την απόφαση επένδυσης. Ο περιορισμός (Narrow) δίνει μια πολύ συμπυκνωμένη εικόνα, αλλά ενδέχεται να παραλείπει σημαντικά στοιχεία που προκύπτουν από τη γενικότερη ανάλυση.

Το παράδειγμα αυτό αναδεικνύει τη σημασία του της βελτιστοποίησης προτροπής (prompt optimization): η ακρίβεια, η σαφήνεια και η κατάλληλη δομή του prompt καθορίζουν την ποιότητα της ανάλυσης και την ευχρηστία της για τον τελικό χρήστη. Επίσης, δείχνει ότι μικρές αλλαγές στη διατύπωση μπορούν να αλλάξουν σημαντικά τον τόνο, τη λεπτομέρεια και την επενδυτική αξία της παραγόμενης σύνοψης.

Συμπερασματικά, η μέθοδος APO (Automatic Prompt Optimization) είναι ένα ισχυρό εργαλείο στο prompt engineering, καθώς επιτρέπει την αξιολόγηση και επιλογή της βέλτιστης διατύπωσης για τη μεγιστοποίηση της ακρίβειας, της πληρότητας και της χρησιμότητας των αποτελεσμάτων.

Συνολικά, η χρήση αυτών των τεχνικών οδήγησε σε σημαντική βελτίωση της ποιότητας, αξιοπιστίας και χρησιμότητας του chatbot. Οι χρήστες μπορούσαν να λάβουν συνοπτικές, κατανοητές και εφαρμόσιμες πληροφορίες για την αγορά, ενώ η εφαρμογή παρέμεινε πλήρως διαδραστική μέσω streamlit. Παρά τα οφέλη αυτά, παραμένουν περιορισμοί λόγω της φύσης των δεδομένων και της αγοράς, καθώς το LLM δεν μπορεί να εγγυηθεί απόλυτη ακρίβεια στις προβλέψεις μετοχών.

Η μελέτη επιβεβαιώνει ότι το prompt engineering αποτελεί βασική δεξιότητα για την αξιοποίηση LLM, ειδικά σε περιβάλλοντα όπου η ακρίβεια, η συνέπεια και η αξιοπιστία είναι κρίσιμες.

## ( 4.5 Συζήτηση Αποτελεσμάτων Red Flagging)

---

Η ενσωμάτωση ενός συστήματος red flagging στη διαχείριση κινδύνων επιτρέπει την αυτόματη ανίχνευση ύποπτων ή επικίνδυνων γεγονότων, παρέχοντας έγκαιρη προειδοποίηση στους υπεύθυνους λήψης αποφάσεων. Το σύστημα αξιοποιεί δεδομένα από διαφορετικές πηγές, όπως ιστορικά τιμών, ειδήσεις ή συναλλαγές, και τα επεξεργάζεται με τεχνικές ανάλυσης και LLM για την αναγνώριση πιθανών κινδύνων ή μη φυσιολογικών συμπεριφορών.

Το παράδειγμα με την ανακοίνωση της Starbucks αναδεικνύει πώς τα μεγάλα γλωσσικά μοντέλα μπορούν να αξιοποιηθούν για την ανίχνευση ασαφούς, υπερβολικά θετικής ή παραπλανητικής γλώσσας σε εταιρικά κείμενα. Η ανακοίνωση περιλαμβάνει φράσεις όπως «τα αποτελέσματα αυτού του τριμήνου δεν αντικατοπτρίζουν τη δύναμη του brand μας», «είμαστε πολύ σίγουροι για το μακροπρόθεσμο μέλλον μας», και «θα αξιοποιήσουμε το απεριόριστο δυναμικό αυτού του brand», οι οποίες δεν προσφέρουν συγκεκριμένα δεδομένα ή μετρήσιμες πληροφορίες και προβάλλουν ένα υπερβολικά θετικό μήνυμα ανεξάρτητα από τα πραγματικά αποτελέσματα.

Η διαδικασία ανάλυσης περιλαμβάνει τον εντοπισμό τριών τύπων γλώσσας: αόριστη/ασαφής (vague), υπερβολικά θετική (overly positive) και παραπλανητική/μη ρεαλιστική (misleading). Στην ανακοίνωση της Starbucks, οι φράσεις αυτές επισημαίνονται ως red flags, υποδεικνύοντας ότι ενώ ο λόγος φαίνεται αισιόδοξος και στρατηγικός, δεν παρέχει ουσιαστικές πληροφορίες για τα οικονομικά αποτελέσματα ή την απόδοση της εταιρείας.

Η χρήση ενός μοντέλου όπως το LLaMA 3 επιτρέπει την συστηματική και γρήγορη αξιολόγηση της γλώσσας, βοηθώντας τους επενδυτές και αναλυτές να ξεχωρίσουν τις δηλώσεις που μπορεί να δημιουργούν θετική εντύπωση χωρίς να υποστηρίζονται από δεδομένα. Επιπλέον, η μέθοδος μπορεί

να χρησιμοποιηθεί για τη δημιουργία πινάκων ελέγχου (dashboards) ή αναφορών επισκόπησης (highlight reports) που εντοπίζουν αυτόματα τις αόριστες ή υπερβολικά αισιόδοξες δηλώσεις σε δεκάδες ανακοινώσεις, αυξάνοντας τη διαφάνεια και τη λογοδοσία.

Η χρήση τεχνικών red flagging έδειξε ότι είναι αποτελεσματική στην αναγνώριση ασυνήθιστων τάσεων, υπερβολικών διακυμάνσεων ή αρνητικών ειδήσεων που μπορεί να επηρεάσουν την αγορά ή μια συγκεκριμένη μετοχή. Οι ειδοποιήσεις παρέχουν τη δυνατότητα στους διαχειριστές κινδύνου να αντιδρούν άμεσα, μειώνοντας πιθανές απώλειες και βελτιώνοντας τη στρατηγική διαχείρισης χαρτοφυλακίων.

Το σύστημα επιτρέπει επίσης τη διαβάθμιση των κινδύνων ανάλογα με τη σοβαρότητα ή τη συχνότητα εμφάνισης των γεγονότων. Για παράδειγμα, μια σειρά αρνητικών ειδήσεων για μια μετοχή μπορεί να ενεργοποιήσει διαφορετικούς δείκτες προτεραιότητας, επιτρέποντας την κατάλληλη κατανομή πόρων για περαιτέρω ανάλυση. Η ενσωμάτωση του red flagging με διαδραστική διεπαφή μέσω streamlit βελτίωσε την ευχρηστία και την προσβασιμότητα των αποτελεσμάτων, επιτρέποντας στους χρήστες να παρακολουθούν σε πραγματικό χρόνο τις προειδοποιήσεις και να λαμβάνουν αποφάσεις με βάση τα δεδομένα.

Παρά τα οφέλη, υπάρχουν περιορισμοί που σχετίζονται κυρίως με την ποιότητα των δεδομένων και τη διαχείριση ψευδώς θετικών (false positives). Ορισμένα γεγονότα μπορεί να ενεργοποιούν red flags χωρίς να αποτελούν πραγματικό κίνδυνο, γεγονός που απαιτεί προσεκτική ρύθμιση των κανόνων και της λογικής του συστήματος. Επίσης, η αξιοπιστία των ειδοποιήσεων εξαρτάται από την ταχύτητα και την ακρίβεια της συλλογής δεδομένων από τις πηγές τους.

## ΚΕΦΑΛΑΙΟ 5 Συμπεράσματα

---

Η παρούσα εργασία ανέδειξε τις δυνατότητες και τους περιορισμούς των Μεγάλων Γλωσσικών Μοντέλων (LLM), και ειδικότερα του Llama 3, στην ανάλυση, επεξεργασία και αξιοποίηση χρηματιστηριακών δεδομένων. Μέσα από την ανάπτυξη τεσσάρων διακριτών εφαρμογών, (α) πρόβλεψης απόδοσης μετοχών, (β) σύνοψης οικονομικών αναφορών, (γ) χρηματιστηριακού chatbot και (δ) συστήματος red-flagging, αξιολογήθηκε σε βάθος τόσο η τεχνική απόδοση όσο και η πρακτική χρησιμότητα των LLM σε πραγματικά επενδυτικά σενάρια.

Η πλατφόρμα streamlit διευκόλυνε τη δημιουργία ενός πλήρως διαδραστικού περιβάλλοντος, ενώ εργαλεία όπως το yfinance και το finvizfinance παρείχαν πηγές αξιόπιστων δεδομένων για περαιτέρω ανάλυση. Η συνεργασία αυτών των εργαλείων με το Llama 3 απέδειξε ότι ο συνδυασμός ποσοτικών (χρονικές σειρές, τεχνικοί δείκτες) και ποιοτικών δεδομένων (ειδήσεις, επιχειρηματικά κείμενα) προσφέρει μια πολύπλευρη προσέγγιση ανάλυσης, απαραίτητη για σύγχρονες χρηματιστηριακές εφαρμογές.

Στην πρόβλεψη απόδοσης μετοχών, τα μοντέλα μηχανικής μάθησης εμφάνισαν χαμηλές τιμές συντελεστή  $R^2$  αλλά καλή κατευθυντική ακρίβεια, η οποία αποτελεί σημαντικό παράγοντα σε περιβάλλοντα υψηλού θορύβου όπως οι χρηματοπιστωτικές αγορές. Το Gradient Boosting αναδείχθηκε το πιο αποτελεσματικό μοντέλο, επιτυγχάνοντας την καλύτερη ισορροπία μεταξύ σφαλμάτων και αξιοπιστίας πρόβλεψης.

Στη σύνοψη οικονομικών αναφορών, το Llama 3 απέδειξε υψηλή ποιότητα αναλυτικών απαντήσεων, ιδιαίτερα στη λεπτομερή εκδοχή του, αν και παρουσίασε σημαντικούς χρόνους επεξεργασίας όταν τρέχει τοπικά σε CPU. Η αριθμητική ακρίβεια κρίθηκε ικανοποιητική για βασικά μεγέθη, ενώ η μετάφραση οικονομικών όρων σε διαφορετικές γλώσσες παραμένει πρόκληση και απαιτεί ανθρώπινη εποπτεία.

Το χρηματιστηριακό chatbot αξιοποίησε τεχνικές prompt engineering (Zero-shot, Few-shot, Chain-of-Thought, ReAct, Prompt Chaining και APO), αποδεικνύοντας ότι η ποιότητα των απαντήσεων ενός LLM εξαρτάται σε μεγάλο βαθμό από τον τρόπο διατύπωσης και δόμησης των prompts. Οι τεχνικές αυτές αύξησαν θεαματικά την πληρότητα, τη σταθερότητα και την επενδυτική αξία των απαντήσεων.

Το σύστημα red flagging έδειξε πώς τα LLM μπορούν να εντοπίσουν ασαφείς, υπερβολικά θετικές ή παραπλανητικές δηλώσεις σε εταιρικά κείμενα, ενισχύοντας την ανίχνευση εταιρικού ρίσκου. Η υλοποίηση με το Llama 3 ανέδειξε την αποτελεσματικότητα των μοντέλων στη γλωσσική αξιολόγηση, ενώ η διαχείριση ψευδώς θετικών (false positives) και η ποικιλία ύφους των εταιρικών ανακοινώσεων αποτελούν σημεία προσοχής.

Εν κατακλείδι, η εργασία αναδεικνύει ότι τα LLM μπορούν να λειτουργήσουν ως ισχυρά εργαλεία υποστήριξης λήψης αποφάσεων στον χρηματοοικονομικό τομέα, προσφέροντας γρήγορη επεξεργασία μεγάλου όγκου δεδομένων, αυξημένη κατανόηση σύνθετων κειμένων και βελτιωμένες αναλυτικές δυνατότητες. Εντούτοις, δεν μπορούν να αντικαταστήσουν πλήρως την ανθρώπινη κρίση, απαιτώντας εποπτεία, επικύρωση και σωστή παραμετροποίηση.

Η ενσωμάτωσή τους σε επενδυτικές εφαρμογές ανοίγει νέες προοπτικές στον τομέα της αυτοματοποίησης και της τεχνητής νοημοσύνης, με την προϋπόθεση ότι θα χρησιμοποιούνται υπεύθυνα, σε συνεργασία με ποσοτικές μεθόδους και εξειδικευμένη ανθρώπινη γνώση.

## BIBΛΙΟΓΡΑΦΙΑ

---

- [1] Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., et al. (2024). Large Language Models: A Survey. *Computer Science, Computation and Language*. arXiv:2402.06196.
- [2] Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., Liu, P., Nie, J.-Y., & Wen, J.-R. (2023). A Survey of Large Language Models. *Computer Science, Computation and Language*. arXiv:2303.18223.
- [3] Matarazzo, A., & Torlone, R. (2025). A Survey on Large Language Models with Some Insights on their Capabilities and Limitations. *arXiv*. arXiv:2501.04040.
- [4] Meta AI. (2024, July). The Llama 3 Herd of Models. *arXiv preprint*. arXiv:2407.21783. <https://doi.org/10.48550/arXiv.2407.21783>
- [5] Wilkinson, L., & Friendly, M. (2009). The History of the Cluster Heat Map. *The American Statistician*, 63, 179–184. <https://doi.org/10.1198/tas.2009.0033>
- [6] Tandon, R. (2024). Prediction of Stock Market Trends Based on Large Language Models. *ITEGAM-Journal of Engineering and Technology for Industrial Applications (ITEGAM-JETIA)*, 11.
- [7] Papatotiriou, K., et al. (2024, October 30). AI in Investment Analysis: LLM for Equity Stock Ratings. *arXiv*. arxiv.org/abs/2411.00856
- [8] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [9] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning Representations by Back-Propagating Errors. *Nature*, 323(6088), 533–536.
- [10] Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely Randomized Trees. *Machine Learning*, 63(1), 3–42. <https://doi.org/10.1007/s10994-006-6226-1>.
- [11] Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support Vector Regression Machines. *Advances in Neural Information Processing Systems*, 9, 155–161.
- [12] Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- [13] Bansal, M., Goyal, A., & Choudhary, A. (2022). Stock Market Prediction with High Accuracy Using Machine Learning Techniques. *Procedia Computer Science*, 215, 247–265. <https://doi.org/10.1016/j.procs.2022.12.028>

- [14] Jin, H., Zhang, Y., Meng, D., Wang, J., & Tan, J. (2024). A Comprehensive Survey on Process-Oriented Automatic Text Summarization with Exploration of LLM-Based Methods. *arXiv*. arXiv:2403.02901.
- [15] Reynolds, L., & McDonell, K. (2021). Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. *arXiv*. arXiv:2102.07350. <https://arxiv.org/abs/2102.07350>