



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΣΧΕΔΙΑΣΜΟΣ ΚΑΙ ΑΝΑΠΤΥΞΗ ANDROID
ΕΦΑΡΜΟΓΗΣ CHATBOT ΓΙΑ ΤΗΝ ΑΝΙΧΝΕΥΣΗ
ΣΥΜΠΤΩΜΑΤΩΝ ΑΓΧΟΥΣ ΚΑΙ ΚΑΤΑΘΛΙΨΗΣ
ΣΕ ΦΟΙΤΗΤΙΚΟ ΠΛΗΘΥΣΜΟ

Μπαλκάμος Σπυρίδων
ΑΜ: 2120203

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΥΠΕΥΘΥΝΟΣ

Ζυγούρης Νικόλαος
Αναπληρωτής Καθηγητής

ΣΥΝΕΠΙΒΛΕΠΩΝ

Κολομβάτσος Κωνσταντίνος
Αναπληρωτής Καθηγητής

Λαμία έτος 2026



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΣΧΕΔΙΑΣΜΟΣ ΚΑΙ ΑΝΑΠΤΥΞΗ ANDROID
ΕΦΑΡΜΟΓΗΣ CHATBOT ΓΙΑ ΤΗΝ ΑΝΙΧΝΕΥΣΗ
ΣΥΜΠΤΩΜΑΤΩΝ ΑΓΧΟΥΣ ΚΑΙ ΚΑΤΑΘΛΙΨΗΣ
ΣΕ ΦΟΙΤΗΤΙΚΟ ΠΛΗΘΥΣΜΟ

Μπαλκάμος Σπυρίδων
ΑΜ: 2120203

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΥΠΕΥΘΥΝΟΣ

Ζυγούρης Νικόλαος
Αναπληρωτής Καθηγητής

ΣΥΝΕΠΙΒΛΕΠΩΝ

Κολομβάτσος Κωνσταντίνος
Αναπληρωτής Καθηγητής

Λαμία έτος 2026



UNIVERSITY OF
THESSALY

SCHOOL OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE & TELECOMMUNICATIONS

DESIGN AND DEVELOPMENT OF AN
ANDROID CHATBOT APPLICATION FOR
DETECTION OF SYMPTOMS OF ANXIETY
AND DEPRESSION IN A STUDENT
POPULATION

Balkamos Spyridon
RN: 2120203

FINAL THESIS

ADVISOR

Zygouris Nikolaos
Associate Professor

CO ADVISOR

Kolomvatsos Konstantinos
Associate Professor

Lamia year 2026

«Με απομική μου ευθύνη και γνωρίζοντας τις κυρώσεις ⁽¹⁾, 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: 25/1/2026

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 παρουσιάζει τον σχεδιασμό και την υλοποίηση ενός πρωτοτύπου συστήματος κινητής υγείας (mHealth) με τη μορφή διαλογικού πράκτορα (chatbot) για συσκευές Android. Το σύστημα στοχεύει στην ανίχνευση ενδείξεων άγχους και κατάθλιψης σε φοιτητές τριτοβάθμιας εκπαίδευσης. Οι αυξανόμενες ακαδημαϊκές πιέσεις καθιστούν αναγκαία την ύπαρξη ψηφιακών εργαλείων προ-διαλογής (screening) που είναι άμεσα προσβάσιμα και πολιτισμικά προσαρμοσμένα. Ωστόσο, η ανάπτυξη τέτοιων συστημάτων αντιμετωπίζει δύο βασικά εμπόδια: την έλλειψη κλινικών δεδομένων λόγω περιορισμών του GDPR και την περιορισμένη υποστήριξη της Ελληνικής γλώσσας στα σύγχρονα μοντέλα Επεξεργασίας Φυσικής Γλώσσας (NLP).

Για την αντιμετώπιση αυτών των προκλήσεων, η εργασία προτείνει μια μεθοδολογία παραγωγής συνθετικών δεδομένων βασισμένη σε στατιστικές κατανομές (Gaussian Personas). Η μεθοδολογία συνδυάζει ψυχομετρικά εργαλεία αξιολόγησης (SCL-90, BDI-II) με τις δυνατότητες παραγωγής κειμένου των Μεγάλων Γλωσσικών Μοντέλων (LLMs), επιτρέποντας τη δημιουργία δεδομένων εκπαίδευσης χωρίς χρήση πραγματικών προσωπικών δεδομένων.

Το σύστημα χρησιμοποιεί υβριδική αρχιτεκτονική συμπερασμού που συνδυάζει δύο στοιχεία: ένα μοντέλο BERT προσαρμοσμένο σε ελληνικά κείμενα για ανάλυση συναισθήματος, και ένα ντετερμινιστικό επίπεδο βασισμένο σε Κανονικές Εκφράσεις (Regex) για τον εντοπισμό κρίσιμων ενδείξεων. Η εργασία παρουσιάζει ένα λειτουργικό πρωτότυπο που αποτελεί τεχνική απόδειξη σκοπιμότητας (proof-of-concept). Απαιτείται μελλοντική επικύρωση σε πραγματικές συνθήκες με δεδομένα πεδίου.

Σημείωση: Το παρόν σύστημα δεν αποτελεί διαγνωστικό εργαλείο και δεν υποκαθιστά την κλινική αξιολόγηση από επαγγελματία ψυχικής υγείας.

ABSTRACT

This thesis presents the design and implementation of a prototype mobile health (mHealth) system: a conversational agent (chatbot) for Android devices that detects indicators of anxiety and depression in university students. Increasing academic pressures have created a need for digital screening tools that are accessible and culturally appropriate. However, developing such systems faces two main obstacles: limited clinical data availability due to GDPR restrictions, and poor support for Modern Greek in current Natural Language Processing (NLP) models.

To address these challenges, this work proposes a synthetic data generation methodology based on statistical distributions (Gaussian Personas). This approach combines psychometric assessment instruments (SCL-90, BDI-II) with the text generation capabilities of Large Language Models (LLMs), enabling the creation of training data without using real personal information.

The system uses a hybrid inference architecture with two components: a BERT model fine-tuned on Greek text for sentiment analysis, and a deterministic layer using Regular Expressions (Regex) to identify critical indicators. The thesis presents a functional prototype that serves as a technical proof-of-concept. Future validation with real-world field data is required.

Note: This system is not a diagnostic tool and does not replace clinical evaluation by a mental health professional.

Table of Contents

ΠΕΡΙΛΗΨΗ	I
ABSTRACT	III
<u>ΚΕΦΑΛΑΙΟ 1 ΕΙΣΑΓΩΓΗ</u>	<u>3</u>
1.1 ΓΕΝΙΚΟ ΠΛΑΙΣΙΟ ΚΑΙ ΚΟΙΝΩΝΙΚΟΟΙΚΟΝΟΜΙΚΗ ΑΝΑΛΥΣΗ	3
1.2 Το ΠΡΟΒΛΗΜΑ: ΑΝΕΠΑΡΚΕΙΑ ΔΟΜΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΟΙ ΠΕΡΙΟΡΙΣΜΟΙ	4
1.3 ΣΚΟΠΟΣ ΚΑΙ ΣΤΟΧΟΙ ΤΗΣ ΔΙΑΤΡΙΒΗΣ.....	5
1.4 ΣΥΝΕΙΣΦΟΡΑ ΤΗΣ ΕΡΓΑΣΙΑΣ	6
<u>ΚΕΦΑΛΑΙΟ 2: ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ.....</u>	<u>7</u>
2.1 ΚΛΙΝΙΚΟ ΠΛΑΙΣΙΟ: ΑΓΧΟΣ ΚΑΙ ΚΑΤΑΘΛΙΨΗ.....	7
2.2 CHATBOTS ΣΤΗΝ ΨΥΧΙΚΗ ΥΓΕΙΑ: ΕΠΙΣΚΟΠΗΣΗ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΙ	7
2.3 ΕΠΕΞΕΡΓΑΣΙΑ ΦΥΣΙΚΗΣ ΓΛΩΣΣΑΣ (NLP) ΚΑΙ ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ.....	8
2.4 ΕΠΕΞΗΓΗΣΙΜΗ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ (ΧΑΙ) ΣΤΗΝ ΨΥΧΙΚΗ ΥΓΕΙΑ.....	9
2.4.A Το ΠΑΡΑΔΟΞΟ ΤΟΥ ΔΙΠΛΟΥ ΜΑΥΡΟΥ ΚΟΥΤΙΟΥ	9
2.4.B ΤΕΧΝΙΚΕΣ ΕΡΜΗΝΕΥΣΙΜΟΤΗΤΑΣ: ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ	10
2.4.Γ ΥΒΡΙΔΙΚΑ ΣΥΣΤΗΜΑΤΑ ΚΑΙ ΔΙΑΦΑΝΕΙΑ	10
2.5 ΠΡΟΣΤΑΣΙΑ ΙΔΙΩΤΙΚΟΤΗΤΑΣ ΚΑΙ GDPR ΣΤΗΝ ΨΗΦΙΑΚΗ ΨΥΧΙΑΤΡΙΚΗ	11
2.5.A ΚΑΝΟΝΙΣΤΙΚΟ ΠΛΑΙΣΙΟ: ΕΥΑΙΣΘΗΤΑ ΔΕΔΟΜΕΝΑ ΚΑΙ ΔΙΚΑΙΩΜΑ ΣΤΗΝ ΕΞΗΓΗΣΗ	11
2.5.B ΤΕΧΝΙΚΕΣ PRIVACY-PRESERVING MACHINE LEARNING (PPML)	11
2.5.Γ ΗΘΙΚΕΣ ΠΡΟΚΛΗΣΕΙΣ ΚΑΙ ΕΜΠΙΣΤΟΣΥΝΗ ΑΣΘΕΝΩΝ	11
2.6 Η ΧΡΗΣΗ ΣΥΝΘΕΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΣΤΗΝ ΙΑΤΡΙΚΗ ΈΡΕΥΝΑ	12
2.7 ΚΡΙΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΚΑΙ ΚΕΝΑ ΣΤΗ ΒΙΒΛΙΟΓΡΑΦΙΑ	12
<u>ΚΕΦΑΛΑΙΟ 3: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ ΚΑΙ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ</u>	<u>15</u>
3.1 ΕΙΣΑΓΩΓΗ ΚΑΙ ΟΡΙΣΜΟΙ.....	15
3.2 ΘΕΜΕΛΙΩΔΕΙΣ ΑΡΧΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	15
3.2.A Ο ΤΕΧΝΗΤΟΣ ΝΕΥΡΩΝΑΣ ΚΑΙ ΤΟ PERCEPTRON.....	15
3.2.B ΣΥΝΑΡΤΗΣΕΙΣ ΕΝΕΡΓΟΠΟΙΗΣΗΣ (ACTIVATION FUNCTIONS).....	15
3.3 ΕΠΕΞΕΡΓΑΣΙΑ ΦΥΣΙΚΗΣ ΓΛΩΣΣΑΣ: ΑΠΟ ΤΑ RNNs ΣΤΟΥΣ TRANSFORMERS	17
3.3.A Η ΜΕΤΑΒΑΣΗ ΑΠΟ ΤΑ RNNs ΣΤΟΥΣ TRANSFORMERS	17
3.3.B ΜΗΧΑΝΙΣΜΟΣ ΑΥΤΟ-ΠΡΟΣΟΧΗΣ (SELF-ATTENTION).....	17
3.3.Γ ΠΟΛΥΚΕΦΑΛΙΚΟΣ ΜΗΧΑΝΙΣΜΟΣ ΠΡΟΣΟΧΗΣ (MULTI-HEAD ATTENTION)	18
3.3.Δ ΚΩΔΙΚΟΠΟΙΗΣΗ ΘΕΣΗΣ (POSITIONAL ENCODINGS).....	18
3.3.Ε ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΥΠΟΛΟΓΙΣΤΙΚΗ ΨΥΧΙΑΤΡΙΚΗ	18
3.4 ΜΟΝΤΕΛΑ BERT ΚΑΙ ΠΡΟΣΑΡΜΟΓΗ	19
3.4.A ΔΙΑΔΙΚΑΣΙΑ FINE-TUNING.....	19
3.5 ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ: ADAMW ΚΑΙ ΚΛΙΝΙΚΕΣ ΜΕΤΡΙΚΕΣ	20
3.5.A ΟΠΙΣΘΟΔΙΑΔΟΣΗ (BACKPROPAGATION).....	20

3.5.Β Ο ΑΛΓΟΡΙΘΜΟΣ ADAMW	20
3.5.Γ ΚΛΙΝΙΚΗ ΑΞΙΟΛΟΓΗΣΗ: MSE/RMSE vs MAE/R ²	21
3.6 ΔΙΑΔΙΚΑΣΙΑ ΕΚΠΑΙΔΕΥΣΗΣ ΚΑΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ	21
3.6.Α ΣΥΝΑΡΤΗΣΕΙΣ ΚΟΣΤΟΥΣ	21
3.6.Β ΟΠΙΣΘΟΔΙΑΔΟΣΗ (BACKPROPAGATION).....	21
3.7 ΔΙΑΔΙΚΑΣΙΑ ΕΚΠΑΙΔΕΥΣΗΣ ΚΑΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ (ΣΥΝΕΧΕΙΑ)	22
3.7.Α ΑΝΑΛΥΣΗ ΚΑΙ ΙΣΟΡΡΟΠΙΑ ΔΕΔΟΜΕΝΩΝ	22
3.7.Β ΕΠΑΥΞΗΣΗ ΔΕΔΟΜΕΝΩΝ (DATA AUGMENTATION).....	23
3.7.Γ ΑΛΓΟΡΙΘΜΟΣ ADAMW ΚΑΙ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ	24
3.7.Δ ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΚΑΙ ΠΑΡΑΚΟΛΟΥΘΗΣΗ ΣΥΓΚΛΙΣΗΣ	24
3.8 ΠΛΑΙΣΙΑ ΑΝΑΠΤΥΞΗΣ ΚΑΙ ΒΙΒΛΙΟΘΗΚΕΣ (FRAMEWORKS & LIBRARIES).....	26

ΚΕΦΑΛΑΙΟ 4: ΜΕΘΟΔΟΛΟΓΙΑ ΥΛΟΠΟΙΗΣΗΣ ΚΑΙ ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΣΥΣΤΗΜΑΤΟΣ..... 27

4.1 ΕΙΣΑΓΩΓΗ ΣΤΗ ΡΟΗ ΕΡΓΑΣΙΑΣ (SYSTEM PIPELINE)	27
4.2 ΜΕΘΟΔΟΛΟΓΙΑ ΠΑΡΑΓΩΓΗΣ ΣΥΝΘΕΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ.....	27
4.2.Α ΣΤΟΧΑΣΤΙΚΗ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΚΛΙΝΙΚΩΝ ΠΡΟΦΙΛ (GAUSSIAN PERSONAS) 28	
4.2.Β ΜΗΧΑΝΙΚΗ ΠΡΟΤΡΟΠΩΝ (PROMPT ENGINEERING)	29
4.3 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗ	30
4.3.Α ΚΑΘΑΡΙΣΜΟΣ ΚΑΙ TOKENIZATION.....	30
4.3.2 TOKENIZATION (ΤΕΜΑΧΙΣΜΟΣ).....	30
4.3.Γ ΣΤΡΑΤΗΓΙΚΗ ΕΚΠΑΙΔΕΥΣΗΣ (MULTI-TASK LEARNING)	31
4.4 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΕΞΥΠΗΡΕΤΗΤΗ (BACKEND IMPLEMENTATION)	31
4.4.Α ΥΒΡΙΔΙΚΗ ΜΗΧΑΝΗ ΣΥΜΠΕΡΑΣΜΟΥ.....	31
4.5 ΑΝΑΠΤΥΞΗ ΕΦΑΡΜΟΓΗΣ ANDROID (FRONTEND)	32
4.5.Α ΕΠΙΚΟΙΝΩΝΙΑ ΜΕ ΤΟ BACKEND	32
4.6 ΔΙΑΣΦΑΛΙΣΗ ΠΟΙΟΤΗΤΑΣ ΚΑΙ GDPR.....	32
4.7 ΔΗΛΩΣΗ ΗΘΙΚΗΣ (ETHICS STATEMENT)	33

ΚΕΦΑΛΑΙΟ 5: ΤΕΧΝΙΚΗ ΥΛΟΠΟΙΗΣΗ..... 34

5.1 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΕΦΑΡΜΟΓΗΣ	34
5.1.Α ΔΥΝΑΜΙΚΗ ΕΠΙΛΟΓΗ ΣΥΣΚΕΥΗΣ	34
5.2 ΥΛΟΠΟΙΗΣΗ BACKEND (FASTAPI & PYTHON).....	35
5.2.Α ΜΗΧΑΝΙΣΜΟΣ ΠΑΡΑΓΩΓΗΣ ΣΥΝΘΕΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ (SYNTHETIC DATA PIPELINE).....	35
5.2.Β ΥΒΡΙΔΙΚΗ ΜΗΧΑΝΗ ΣΥΜΠΕΡΑΣΜΟΥ: ΣΥΝΤΗΞΗ BERT ΚΑΙ ΕΥΡΕΤΙΚΩΝ ΚΑΝΟΝΩΝ (HYBRID FUSION).....	37
5.2.Γ ΕΚΠΑΙΔΕΥΣΗ ΜΟΝΤΕΛΟΥ ΚΑΙ ΤΕΧΝΙΚΕΣ ΕΠΑΥΞΗΣΗΣ ΔΕΔΟΜΕΝΩΝ	45
5.3 ΕΦΑΡΜΟΓΗ ANDROID (CLIENT).....	45
5.3.Α ΒΑΣΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ:.....	46
5.3.Β ΕΝΗΜΕΡΩΜΕΝΗ ΣΥΓΚΑΤΑΘΕΣΗ	47
5.3.Γ ΠΟΛΙΤΙΚΗ ΑΣΦΑΛΕΙΑΣ ΔΙΚΤΥΟΥ	47
5.3.Δ ΣΥΛΛΟΓΗ ΔΗΜΟΓΡΑΦΙΚΩΝ.....	48
5.3.Ε DATA TRANSFER OBJECTS ΚΑΙ MAPPER PATTERN	49
5.3.ΣΤ ΣΥΛΛΟΓΗ ΕΛΕΥΘΕΡΟΥ ΚΕΙΜΕΝΟΥ.....	50
5.3.Ζ ΑΣΦΑΛΗΣ ΑΠΟΣΤΟΛΗ ΔΕΔΟΜΕΝΩΝ	51
5.3.Η ΨΥΧΟΜΕΤΡΙΚΑ ΕΡΩΤΗΜΑΤΟΛΟΓΙΑ.....	52
5.3.Θ ΠΡΟΤΑΣΕΙΣ ΓΙΑ PRODUCTION-GRADE SECURITY	54

5.3.Ι ΟΘΟΝΗ ΕΠΙΣΚΟΠΗΣΗΣ.....	55
5.3.Κ ΔΙΑΧΕΙΡΙΣΗ ΑΣΥΓΧΡΟΝΗΣ ΕΠΙΚΟΙΝΩΝΙΑΣ ΜΕ MVVM ΚΑΙ COROUTINES.....	56
5.3.Λ ΟΘΟΝΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ	56
5.3.Μ ΑΣΥΓΧΡΟΝΗ ΑΝΑΛΥΣΗ	57
5.4 ΠΡΟΚΛΗΣΕΙΣ ΥΛΟΠΟΙΗΣΗΣ ΚΑΙ ΣΤΡΑΤΗΓΙΚΕΣ ΑΝΤΙΜΕΤΩΠΙΣΗΣ.....	58

ΚΕΦΑΛΑΙΟ 6: ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ 60

6.1 ΡΥΘΜΙΣΕΙΣ ΕΚΠΑΙΔΕΥΣΗΣ.....	60
6.2 ΠΟΣΟΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ.....	60
6.2.Α ΕΡΜΗΝΕΙΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ	60
6.3 ΜΕΛΕΤΗ ΕΠΙΔΡΑΣΗΣ ΣΥΝΙΣΤΩΣΩΝ (ABLATION STUDY)	61
6.4 ΠΟΙΟΤΙΚΗ ΑΝΑΛΥΣΗ.....	62
6.4.Α ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ.....	62
6.5 ΧΡΟΝΟΙ ΑΠΟΚΡΙΣΗΣ (LATENCY BENCHMARKING)	67

ΚΕΦΑΛΑΙΟ 7: ΣΥΖΗΤΗΣΗ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΙ..... 68

7.1 ΣΥΓΚΡΙΣΗ ΜΕ ΆΛΛΑ ΣΥΣΤΗΜΑΤΑ.....	68
7.2 ΗΘΙΚΟΙ ΠΡΟΒΛΗΜΑΤΙΣΜΟΙ ΚΑΙ ΕΥΘΥΝΗ	68
7.2.Α ΜΕΡΟΛΗΨΙΑ ΔΕΔΟΜΕΝΩΝ (DATA BIAS)	68
7.2.Β ΝΟΜΙΚΗ ΕΥΘΥΝΗ (MEDICAL LIABILITY)	69
7.2.Γ ΔΙΑΦΑΝΕΙΑ ΚΑΙ ΕΝΗΜΕΡΗ ΣΥΓΚΑΤΑΘΕΣΗ	69
7.3 ΠΕΡΙΟΡΙΣΜΟΙ ΤΗΣ ΈΡΕΥΝΑΣ	69
7.3.Α ΣΥΝΘΕΤΙΚΑ VS ΠΡΑΓΜΑΤΙΚΑ ΔΕΔΟΜΕΝΑ.....	69
7.3.Β ΈΛΛΕΙΨΗ ΚΛΙΝΙΚΗΣ ΕΠΙΚΥΡΩΣΗΣ	69
7.3.Γ ΜΟΝΟΓΛΩΣΣΙΚΟ ΜΟΝΤΕΛΟ.....	70
7.3.Δ ΑΠΟΥΣΙΑ ΧΡΟΝΙΚΗΣ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ	70
7.4 ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ (FUTURE WORK).....	70
7.4.Α ΜΕΤΑΒΑΣΗ ΣΕ ΔΕΔΟΜΕΝΑ ΠΡΑΓΜΑΤΙΚΗΣ ΧΡΗΣΗΣ.....	70
7.4.Β MULTIMODAL FEATURES & ΦΩΝΗΤΙΚΗ ΑΝΑΛΥΣΗ	71
7.4.Γ EDGE AI & ON-DEVICE INFERENCE	71

ΚΕΦΑΛΑΙΟ 8: ΣΥΜΠΕΡΑΣΜΑΤΑ..... 72

8.1 ΚΥΡΙΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ	72
8.2 ΠΕΡΙΟΡΙΣΜΟΙ ΚΑΙ ΕΠΙΣΗΜΑΝΣΕΙΣ	72
8.3 ΠΡΟΟΠΤΙΚΗ ΣΥΝΕΙΣΦΟΡΑ	72
8.4 ΑΠΟ ΤΟ PROOF-OF-CONCEPT ΣΤΗΝ ΠΙΘΑΝΗ ΚΛΙΝΙΚΗ ΕΦΑΡΜΟΓΗ.....	73
8.5 ΤΕΛΙΚΟ ΜΗΝΥΜΑ	73

ΒΙΒΛΙΟΓΡΑΦΙΑ..... 74

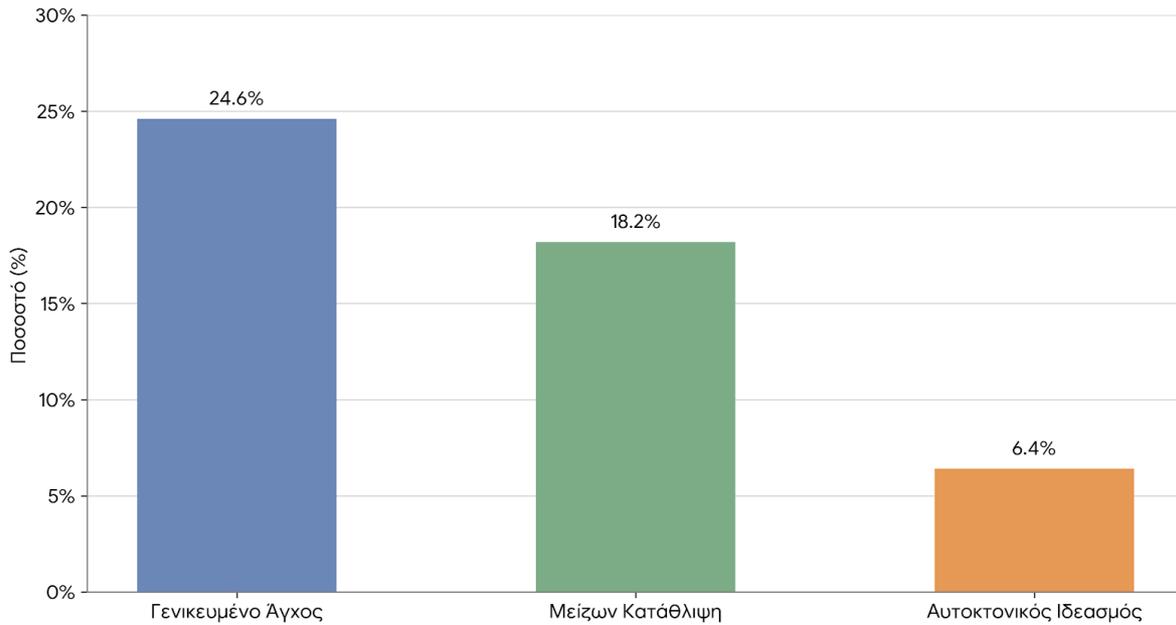
ΚΕΦΑΛΑΙΟ 1 ΕΙΣΑΓΩΓΗ

1.1 Γενικό Πλαίσιο και Κοινωνικοοικονομική Ανάλυση

Η ψυχική υγεία των φοιτητών έχει αναδειχθεί σε σημαντικό ζήτημα δημόσιας υγείας σε παγκόσμια κλίμακα, με τα ποσοστά των διαταραχών άγχους και κατάθλιψης να παρουσιάζουν σημαντική αύξηση κατά την τελευταία δεκαετία. Η μετάβαση στην τριτοβάθμια εκπαίδευση αποτελεί μια κρίσιμη αναπτυξιακή περίοδο που χαρακτηρίζεται από την ανάληψη νέων ευθυνών, την απομάκρυνση από το υποστηρικτικό οικογενειακό περιβάλλον, την οικονομική διαχείριση και την ακαδημαϊκή πίεση. Οι παράγοντες αυτοί, σε συνδυασμό με την αβεβαιότητα για την επαγγελματική αποκατάσταση, λειτουργούν κάποιες φορές συσσωρευτικά ως ελκτικοί μηχανισμοί για την εμφάνιση ψυχοπαθολογίας [7].

Στην Ελλάδα, η κατάσταση επιδεινώθηκε κατά την περίοδο της πανδημίας COVID-19 και τη μετα-πανδημική εποχή (2023-2024). Επιδημιολογικές μελέτες σε ελληνικά πανεπιστημιακά ιδρύματα, όπως το Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης και το Δημοκρίτειο Πανεπιστήμιο Θράκης, καταγράφουν ανησυχητικά ευρήματα. Η συχνότητα εμφάνισης συμπτωμάτων σοβαρού άγχους και κατάθλιψης φαίνεται να έχει αυξηθεί σημαντικά σε σύγκριση με την προ-πανδημική περίοδο. Ερευνητικά δεδομένα υποδεικνύουν ότι το ποσοστό των φοιτητών που αναφέρουν κλινικά σημαντικά επίπεδα γενικευμένου άγχους ανέρχεται στο 24.6%, ενώ τα ποσοστά της μείζονος κατάθλιψης αγγίζουν το 18.2% του φοιτητικού πληθυσμού [6]. Η συννοσηρότητα μεταξύ άγχους και κατάθλιψης είναι συχνή, επιβαρύνοντας περαιτέρω τη λειτουργικότητα των ατόμων.

Μια ιδιαίτερα κρίσιμη παράμετρος είναι ο αυτοκτονικός ιδεασμός, ο οποίος ανιχνεύεται σε ποσοστό περίπου 6.4% των φοιτητών, υπογραμμίζοντας την ανάγκη για μηχανισμούς έγκαιρης παρέμβασης [1]. Παρά την υψηλή συχνότητα εμφάνισης των διαταραχών, το ποσοστό των φοιτητών που αναζητούν ενεργά επαγγελματική βοήθεια παραμένει χαμηλό, κυμαινόμενο κάτω από το 30% [7]. Οι φραγμοί στην αναζήτηση βοήθειας είναι πολυδιάστατοι και περιλαμβάνουν το κοινωνικό στίγμα που εξακολουθεί να περιβάλλει την ψυχική νόσο στην ελληνική κοινωνία, το κόστος των ιδιωτικών υπηρεσιών ψυχικής υγείας, καθώς και την έλλειψη ενημέρωσης σχετικά με τις διαθέσιμες δομές [11].



Εικόνα 1.1: Ποσοστά διαταραχών ψυχικής υγείας σε φοιτητικούς πληθυσμούς. (Βάσει των ερευνών [6], [7]).

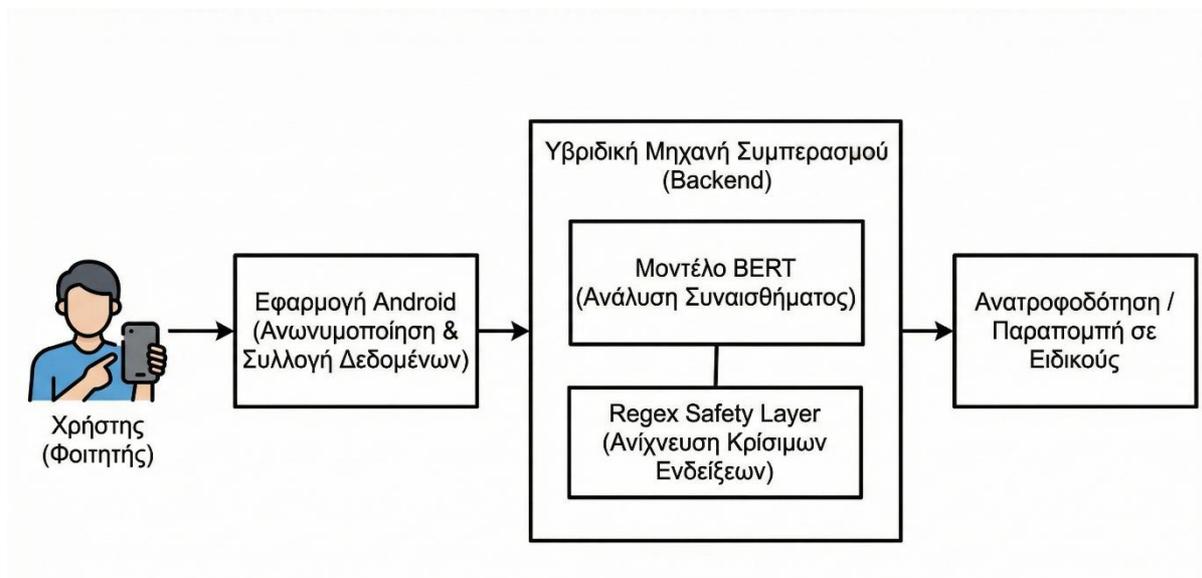
1.2 Το Πρόβλημα: Ανεπάρκεια Δομών και Τεχνολογικοί Περιορισμοί

Η υφιστάμενη κατάσταση χαρακτηρίζεται από ένα σημαντικό «κενό θεραπείας» (treatment gap). Οι συμβουλευτικοί σταθμοί των ελληνικών πανεπιστημίων, αν και στελεχωμένοι με εξειδικευμένο προσωπικό, δυσκολεύονται να ανταποκριθούν στον αυξανόμενο όγκο των αιτημάτων, οδηγώντας σε λίστες αναμονής που μπορεί να διαρκέσουν μήνες [11]. Αυτή η χρονική υστέρηση είναι κρίσιμη, καθώς η έγκαιρη παρέμβαση στα πρώιμα στάδια των συμπτωμάτων είναι καθοριστική για την πρόγνωση.

Επιπροσθέτως, τα ψηφιακά εργαλεία που διατίθενται σήμερα παρουσιάζουν σημαντικούς λειτουργικούς και μεθοδολογικούς περιορισμούς. Τα περισσότερα εργαλεία προ-διαλογής (screening tools) αποτελούν απλές ψηφιοποιήσεις ερωτηματολογίων πολλαπλής επιλογής, χωρίς να επιτρέπουν την ελεύθερη έκφραση του χρήστη και χωρίς να μπορούν να συλλάβουν τις λεπτές αποχρώσεις του συναισθηματικού λόγου που συχνά αποκαλύπτονται μέσω της φυσικής γλώσσας. Παράλληλα, τα συστήματα που βασίζονται σε Τεχνητή Νοημοσύνη, όπως τα Woebot και Wysa, έχουν αναπτυχθεί κυρίως για την Αγγλική γλώσσα. Η εφαρμογή τους σε ελληνόφωνους πληθυσμούς είναι προβληματική, καθώς τα μοντέλα NLP (Natural Language Processing) δυσκολεύονται να διαχειριστούν τη μορφολογική πολυπλοκότητα της Ελληνικής γλώσσας και τα πολιτισμικά συμφραζόμενα που επηρεάζουν τον τρόπο έκφρασης του ψυχικού πόνου [11]. Τέλος, η πλειονότητα των εφαρμογών mHealth βασίζεται σε αρχιτεκτονικές cloud που αποθηκεύουν ευαίσθητα ιατρικά δεδομένα σε διακομιστές τρίτων, συχνά εκτός της Ευρωπαϊκής Ένωσης, εγείροντας ζητήματα συμμόρφωσης με τον Γενικό Κανονισμό Προστασίας Δεδομένων (GDPR) [4].

1.3 Σκοπός και Στόχοι της Διατριβής

Ο κύριος σκοπός της παρούσας εργασίας είναι ο σχεδιασμός και η υλοποίηση ενός πρωτότυπου συστήματος mHealth που αξιοποιεί τις δυνατότητες της Τεχνητής Νοημοσύνης για την ανίχνευση ενδείξεων άγχους και κατάθλιψης μέσω της ανάλυσης φυσικής γλώσσας. Το προτεινόμενο σύστημα στοχεύει να λειτουργήσει ως ένα πρώτο επίπεδο προ-διαλογής (triage), παρέχοντας στον χρήστη ανατροφοδότηση και προτείνοντας αναζήτηση ειδικής βοήθειας όταν εντοπίζονται ενδείξεις υψηλού κινδύνου. Η γενική αρχιτεκτονική του προτεινόμενου συστήματος απεικονίζεται στην Εικόνα 1.2.



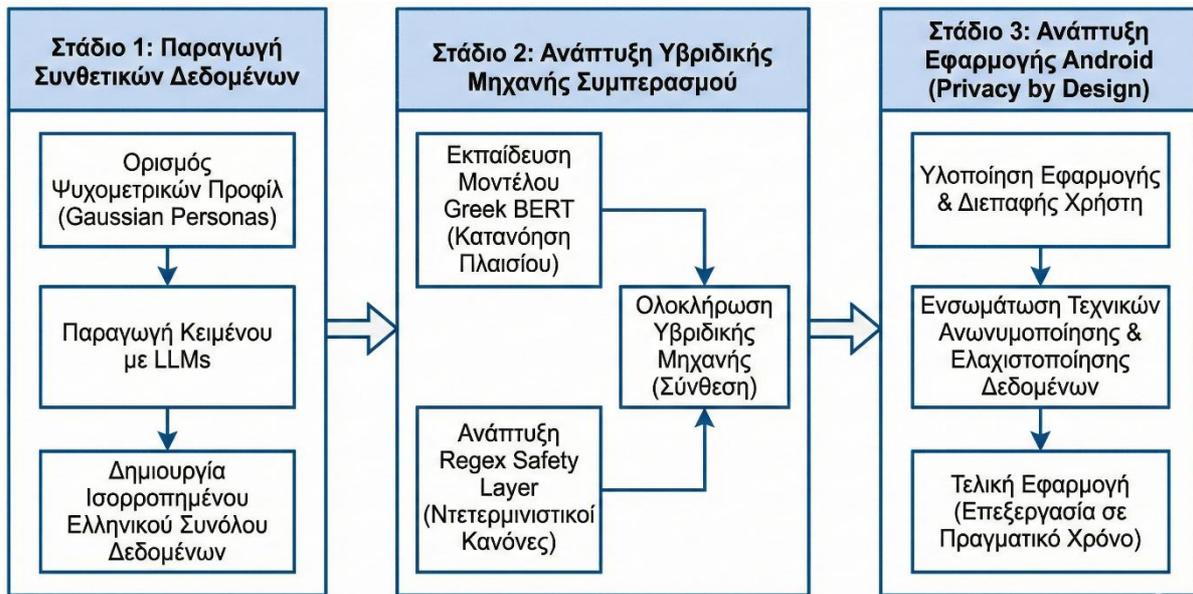
Εικόνα 1.2: Αρχιτεκτονική του προτεινόμενου συστήματος mHealth.

Οι ειδικοί στόχοι της εργασίας εξειδικεύονται ως εξής. Ο πρώτος στόχος αφορά την ανάπτυξη μιας μεθοδολογίας παραγωγής τεχνητών δεδομένων (Synthetic Data Pipeline) που συνδυάζει ψυχομετρικά προφίλ μέσω στατιστικών κατανομών (Gaussian Personas) με τις δυνατότητες παραγωγής κειμένου των Μεγάλων Γλωσσικών Μοντέλων (LLMs). Στόχος είναι η παραγωγή ενός ισορροπημένου συνόλου δεδομένων στην Ελληνική γλώσσα, προτεινόμενου ως λύση στο πρόβλημα της έλλειψης δεδομένων λόγω GDPR [23].

Ο δεύτερος στόχος αφορά τον σχεδιασμό και την υλοποίηση μιας υβριδικής μηχανής συμπερασμού. Η αρχιτεκτονική αυτή συνθέτει την ικανότητα κατανόησης πλαισίου του μοντέλου Greek BERT με ένα ντετερμινιστικό επίπεδο ασφαλείας βασισμένο σε Κανονικές Εκφράσεις (Regex). Στόχος του σχεδιασμού είναι η βελτίωση της ανίχνευσης κρίσιμων περιστατικών μέσω ενός ευρετικού μηχανισμού ασφαλείας (heuristic safety layer).

Ο τρίτος στόχος αφορά την ανάπτυξη εφαρμογής με επίκεντρο την ιδιωτικότητα (Privacy by Design). Υλοποιήθηκε εφαρμογή Android που ελαχιστοποιεί τη συλλογή δεδομένων, χρησιμοποιεί τεχνικές ανωνυμοποίησης και επεξεργάζεται τα δεδομένα σε πραγματικό χρόνο χωρίς μόνιμη αποθήκευση, επιδιώκοντας συμμόρφωση με το άρθρο 4 του GDPR [4].

Η ροή εργασιών για την επίτευξη των παραπάνω στόχων, από την παραγωγή δεδομένων έως την εφαρμογή, παρουσιάζεται στην Εικόνα 1.3.



Εικόνα 1.3: Ροή εργασιών ανάπτυξης του συστήματος.

1.4 Συνεισφορά της Εργασίας

Η παρούσα εργασία συνεισφέρει στο πεδίο της Βιοϊατρικής Πληροφορικής και της Υπολογιστικής Γλωσσολογίας σε τρεις διακριτούς άξονες.

Ο πρώτος άξονας αφορά τη μεθοδολογία συνθετικών/τεχνητών δεδομένων. Παρουσιάζεται μια προσέγγιση που συνδυάζει Μεγάλα Γλωσσικά Μοντέλα με στατιστικές κατανομές για την παραγωγή ψυχομετρικών δεδομένων εκπαίδευσης, εφαρμόζοντας υφιστάμενες μεθοδολογίες στο πλαίσιο της Ελληνικής γλώσσας. Η μέθοδος αυτή καταδεικνύει τη δυνατότητα εκπαίδευσης μοντέλων NLP χωρίς την ανάγκη έκθεσης πραγματικών ασθενών [23].

Ο δεύτερος άξονας εστιάζει στην ασφάλεια του συστήματος μέσω μιας υβριδικής αρχιτεκτονικής. Επειδή τα μοντέλα Τεχνητής Νοημοσύνης περιέχουν πάντα έναν βαθμό αβεβαιότητας, προτείνεται η προσθήκη ενός συστήματος αυστηρών κανόνων που λειτουργεί ως δικλείδα ασφαλείας (fallback mechanism). Στόχος είναι να διασφαλιστεί ότι τα κρίσιμα περιστατικά θα εντοπίζονται πάντα, μεγιστοποιώντας την ευαισθησία του συστήματος σε περιπτώσεις υψηλού κινδύνου.

Ο τρίτος άξονας αφορά την εξειδίκευση στην Ελληνική γλώσσα. Αναπτύχθηκε ένα σύστημα προ-διαλογής ανοικτού κώδικα για την Ελληνική γλώσσα, ευθυγραμμισμένο με τις κλίμακες BDI-II και SCL-90, συμβάλλοντας στην κάλυψη ενός ερευνητικού κενού για τις γλώσσες με περιορισμένους πόρους (low-resource languages) [3]. Θα πρέπει να σημειωθεί ότι, εξ όσων γνωρίζουμε, δεν εντοπίστηκε αντίστοιχο δημοσίως διαθέσιμο σύστημα για την Ελληνική γλώσσα κατά τη βιβλιογραφική επισκόπηση της παρούσας εργασίας.

Τέλος, θα πρέπει να σημειωθεί ότι η παρούσα εργασία αποτελεί, σε επίπεδο αρχιτεκτονικής σχεδίασης, επέκταση της μεθοδολογίας που αναπτύχθηκε από τον Βιολέντη [30], προσαρμόζοντας τα τεχνικά πρότυπα της μηχανικής μάθησης από το πεδίο της ανίχνευσης διαταραχών γραφής στο πεδίο της Ψυχικής Υγείας και της Επεξεργασίας Φυσικής Γλώσσας.

Κεφάλαιο 2: ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ

2.1 Κλινικό Πλαίσιο: Άγχος και Κατάθλιψη

Πριν την ανάλυση των τεχνολογικών μέσων παρέμβασης, κρίνεται απαραίτητη η οριοθέτηση των κλινικών οντοτήτων που καλείται να διαχειριστεί το προτεινόμενο σύστημα. Σύμφωνα με το Διαγνωστικό και Στατιστικό Εγχειρίδιο Ψυχικών Διαταραχών (DSM-5), οι Διαταραχές Άγχους και η Μείζων Καταθλιπτική Διαταραχή αποτελούν δύο από τις συχνότερες ψυχοπαθολογικές καταστάσεις [6].

Η Μείζων Κατάθλιψη χαρακτηρίζεται από επίμονο αίσθημα θλίψης, απώλεια ενδιαφέροντος (ανηδονία), διαταραχές ύπνου και όρεξης, καθώς και γνωστικά ελλείμματα στη συγκέντρωση [11]. Για την ποσοτικοποίηση της σοβαρότητας των συμπτωμάτων, η κλινική πρακτική αξιοποιεί σταθμισμένα ψυχομετρικά εργαλεία. Στην παρούσα εργασία, υιοθετείται το Beck Depression Inventory (BDI-II), ένα ερωτηματολόγιο 21 ερωτήσεων που αξιολογεί τη σωματοποίηση και το συναισθηματικό φορτίο της νόσου [14].

Αντίστοιχα, το Άγχος ορίζεται ως μια δυσάρεστη συναισθηματική κατάσταση που συνοδεύεται από νευρικότητα και σωματικές εκδηλώσεις, όπως ταχυκαρδία και εφίδρωση. Για την αξιολόγησή του, χρησιμοποιείται συχνά η κλίμακα SCL-90 (Symptom Checklist-90), η οποία επιτρέπει την πολυδιάστατη καταγραφή ψυχολογικών συμπτωμάτων [14]. Η κατανόηση αυτών των κλινικών δομών είναι σημαντική για την εκπαίδευση των αλγορίθμων Μηχανικής Μάθησης που θα περιγραφούν στη συνέχεια, καθώς το μοντέλο καλείται να εντοπίσει γλωσσικά μοτίβα που συσχετίζονται με τα κριτήρια του DSM-5.

2.2 Chatbots στην Ψυχική Υγεία: Επισκόπηση και Περιορισμοί

Η ενσωμάτωση των διαλογικών πρακτόρων (chatbots) στον τομέα της ψυχικής υγείας έχει αποτελέσει αντικείμενο εκτεταμένης έρευνας. Ο όρος «ψηφιακές θεραπευτικές παρεμβάσεις» (digital therapeutics) χρησιμοποιείται για να περιγράψει αυτά τα συστήματα, αν και θα πρέπει να σημειωθεί ότι η πλειονότητα δεν έχει λάβει έγκριση ως ιατροτεχνολογικό προϊόν.

Το Woebot αποτελεί ένα από τα πιο διαδεδομένα παραδείγματα διεθνώς. Πρόκειται για σύστημα που βασίζεται στις αρχές της Γνωσιακής Συμπεριφορικής Θεραπείας (CBT) και χρησιμοποιεί κυρίως δέντρα αποφάσεων (decision trees) και προκαθορισμένα σενάρια διαλόγου για να καθοδηγήσει τον χρήστη στην αναδόμηση αρνητικών σκέψεων [14]. Αν και κλινικές δοκιμές έχουν δείξει θετική επίδραση στη μείωση των συμπτωμάτων κατάθλιψης, η λειτουργία του είναι αυστηρά δομημένη (rule-based), περιορίζοντας τη φυσικότητα της αλληλεπίδρασης. Επιπλέον, η έλλειψη υποστήριξης της Ελληνικής γλώσσας καθιστά την εφαρμογή του απαγορευτική για τον πληθυσμό στόχο της παρούσας μελέτης [10].

Το Wysa αποτελεί ένα άλλο σύστημα που υιοθετεί μια πιο ευέλικτη προσέγγιση τεχνητής νοημοσύνης για την παροχή υποστήριξης. Ωστόσο, η αποτελεσματικότητά του βασίζεται κυρίως σε υποκειμενικές αυτο-αναφορές χρηστών και στερείται εργαλείων αντικειμενικής αξιολόγησης σε πραγματικό χρόνο, όπως η εξαγωγή σκορ για τις κλίμακες BDI ή SCL-90 [17]. Επιπλέον, το επιχειρηματικό μοντέλο συνδρομής που ακολουθεί δημιουργεί οικονομικούς φραγμούς πρόσβασης, ειδικά για φοιτητές.

2.3 Επεξεργασία Φυσικής Γλώσσας (NLP) και Ανάλυση Συναισθήματος

Η πρόοδος στην Επεξεργασία Φυσικής Γλώσσας (NLP) έχει επιτρέψει τη μετάβαση από τα στατικά ερωτηματολόγια στη δυναμική ανάλυση ελεύθερου κειμένου. Η σύγχρονη βιβλιογραφία υποδεικνύει ότι μοντέλα βαθιάς μάθησης (Deep Learning), και ειδικότερα η αρχιτεκτονική Transformer στην οποία βασίζεται το BERT, μπορούν να εντοπίσουν γλωσσικά μοτίβα που σχετίζονται με την κατάθλιψη με μεγαλύτερη ακρίβεια από τις παραδοσιακές μεθόδους μηχανικής μάθησης [3]. Σε αντίθεση με παλαιότερες προσεγγίσεις που αντιμετώπιζαν τις λέξεις μεμονωμένα, όπως η μέθοδος Bag-of-Words, οι Transformers χρησιμοποιούν τον μηχανισμό της Αυτο-Προσοχής (Self-Attention). Αυτό επιτρέπει στο μοντέλο να κατανοεί το συμφραζόμενο (context), σταθμίζοντας τη σημασία κάθε λέξης σε σχέση με όλες τις υπόλοιπες λέξεις της πρότασης. Έτσι, το μοντέλο μπορεί να διακρίνει τη διαφορά μεταξύ του «είμαι μόνος στο σπίτι» (ουδέτερο) και του «νιώθω τόσο μόνος» (συναισθηματικό φορτίο).

Στο πεδίο της Ελληνικής γλώσσας, η δημοσίευση του Greek BERT από τους Koutsikakis et al. (2020) αποτέλεσε σημαντική εξέλιξη, παρέχοντας ένα προ-εκπαιδευμένο μοντέλο υψηλής απόδοσης [18]. Η εφαρμογή τέτοιων μοντέλων στα Ελληνικά παρουσιάζει ιδιαίτερες προκλήσεις λόγω της πλούσιας μορφολογίας της γλώσσας, η οποία περιλαμβάνει κλίσεις, καταλήξεις και τόνους. Ένα χαρακτηριστικό πρόβλημα είναι ότι μια λέξη μπορεί να εμφανιστεί με πολλές διαφορετικές μορφές, όπως «γράφω», «έγραψα», «γραμμένο» και «γράφοντας», δυσκολεύοντας το μοντέλο να αντιληφθεί ότι πρόκειται για την ίδια σημασιολογική ρίζα.

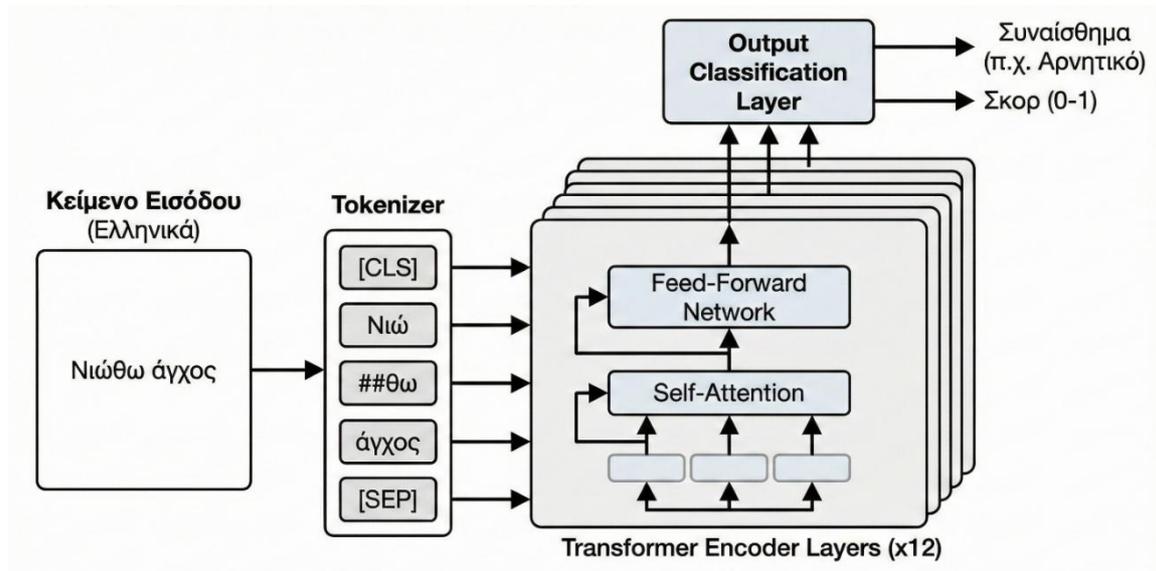
Το Greek BERT αντιμετωπίζει αυτό το πρόβλημα μέσω μιας διαδικασίας που ονομάζεται Τεμαχισμός σε υπο-λέξεις (Sub-word Tokenization), χρησιμοποιώντας τον αλγόριθμο WordPiece [18]. Το μοντέλο δεν επεξεργάζεται ολόκληρες λέξεις, αλλά μαθαίνει να αναγνωρίζει ρίζες και συχνές καταλήξεις. Όταν συναντά μια σύνθετη ή άγνωστη λέξη, τη διασπά σε μικρότερα, γνωστά κομμάτια. Αυτό είναι χρήσιμο για την ανάλυση κειμένων που εκφράζουν ψυχική δυσφορία, όπου οι χρήστες συχνά χρησιμοποιούν σύνθετες εκφράσεις για να περιγράψουν τα συναισθήματά τους. Ένα παράδειγμα αυτής της διαδικασίας παρουσιάζεται στην Εικόνα 2.1.

Βήμα	Δεδομένα	Επεξήγηση
Input	Νιώθω απελπισία	Η αρχική φράση εισόδου.
Tokens	['[CLS]', 'Νιώ', '##θω', 'απελ', '##πισί', '##α', '[SEP]']	Η λέξη "Νιώθω" σπάει σε ρίζα και κατάληξη. Το ## δηλώνει συνέχεια λέξης.
Input IDs	[101, 1543, 2309, 5671, 8901, 1022, 102]	Οι αριθμητικές τιμές που αντιστοιχούν στα tokens και εισάγονται στο νευρωνικό δίκτυο.

Εικόνα 2.1: Παράδειγμα Tokenizer του Greek BERT

Όπως παρατηρείται, οι λέξεις διασπώνται σε σημασιολογικά κομμάτια (tokens), τα οποία στη συνέχεια μετατρέπονται σε μοναδικούς αριθμητικούς κωδικούς (Input IDs). Αυτοί οι κωδικοί αποτελούν την είσοδο στην κύρια αρχιτεκτονική του BERT.

Η συνολική διαδικασία επεξεργασίας, από την εισαγωγή του ακατέργαστου ελληνικού κειμένου μέχρι την τελική πρόβλεψη του μοντέλου, απεικονίζεται στην Εικόνα 2.2. Το σύστημα αξιοποιεί 12 επίπεδα κωδικοποίησης (Transformer Encoder Layers) για να εξάγει σημασιολογικά χαρακτηριστικά από τα tokens, πριν καταλήξει σε μια πρόβλεψη.



Εικόνα 2.2: Διάγραμμα Αρχιτεκτονικής Επεξεργασίας Greek BERT

Παρά τις δυνατότητές του, οι περισσότερες υφιστάμενες εφαρμογές που αξιοποιούν το Greek BERT περιορίζονται σε απλή ταξινόμηση συναίσθηματος, όπως η διάκριση μεταξύ θετικού και αρνητικού, και δεν επεκτείνονται στην αξιολόγηση σοβαρότητας συμπτωμάτων μέσω παλινδρόμησης (regression). Αυτό αποτελεί ένα ερευνητικό κενό το οποίο επιχειρεί να καλύψει η παρούσα εργασία [22].

2.4 Επεξηγήσιμη Τεχνητή Νοημοσύνη (XAI) στην Ψυχική Υγεία

Η ενσωμάτωση αλγορίθμων Βαθιάς Μάθησης (Deep Learning) στην ιατρική πρακτική αντιμετωπίζει ένα σημαντικό εμπόδιο, γνωστό στη βιβλιογραφία ως το Πρόβλημα του Μαύρου Κουτιού (Black Box Problem). Ενώ μοντέλα όπως το BERT επιτυγχάνουν υψηλή ακρίβεια πρόβλεψης, η εσωτερική τους λειτουργία, η οποία βασίζεται σε εκατομμύρια παραμέτρους και μη γραμμικούς μετασχηματισμούς, παραμένει δυσνόητη [46]. Στην ψυχική υγεία, αυτή η αδιαφάνεια δημιουργεί ζητήματα εμπιστοσύνης και ασφάλειας, καθιστώντας την Επεξηγήσιμη Τεχνητή Νοημοσύνη (Explainable AI - XAI) μια σημαντική ερευνητική κατεύθυνση [47].

2.4.α Το Παράδοξο του Διπλού Μαύρου Κουτιού

Η εφαρμογή της TN στην ψυχιατρική διαφέρει από άλλους ιατρικούς κλάδους λόγω του «Παραδόξου του Διπλού Μαύρου Κουτιού» [49]. Το πρώτο «μαύρο κουτί» είναι ο ίδιος ο ανθρώπινος εγκέφαλος και η παθοφυσιολογία των ψυχικών διαταραχών, η οποία συχνά

στερείται αντικειμενικών βιοδεικτών (ground truth) και βασίζεται σε υποκειμενική κλινική συναίνεση. Όταν ένα Βαθύ Νευρωνικό Δίκτυο (το δεύτερο «μαύρο κουτί») καλείται να αναλύσει δεδομένα από το πρώτο, υπάρχει ο κίνδυνος το μοντέλο να εντοπίσει στατιστικές συσχετίσεις χωρίς ουσιαστική αξία (spurious correlations) [49]. Συνεπώς, η ερμηνευσιμότητα είναι σημαντική για να διασφαλιστεί ότι το μοντέλο βασίζεται σε ψυχοπαθολογικά σχετικά χαρακτηριστικά.

2.4.β Τεχνικές Ερμηνευσιμότητας: Θεωρητικό Πλαίσιο

Η βιβλιογραφία της περιόδου 2020-2025 αναδεικνύει τρεις κύριες προσεγγίσεις για τη βελτίωση της ερμηνευσιμότητας των μοντέλων βαθιάς μάθησης.

Η μέθοδος SHAP (SHapley Additive exPlanations) βασίζεται στη συνεργατική θεωρία παιγνίων και χρησιμοποιείται για την ανάλυση της συμβολής κάθε χαρακτηριστικού (feature importance) στην τελική πρόβλεψη. Μελέτες έχουν δείξει ότι το SHAP μπορεί να αποκαλύψει σχετικές παραμέτρους, όπως διαταραχές στον κερκάδιο ρυθμό κατά την ανάλυση δεδομένων δραστηριογράφων για κατάθλιψη [48].

Η μέθοδος LIME (Local Interpretable Model-agnostic Explanations) εστιάζει στην τοπική ερμηνευσιμότητα, δημιουργώντας απλά γραμμικά μοντέλα γύρω από μια συγκεκριμένη πρόβλεψη. Αν και χρήσιμη για την επεξήγηση μεμονωμένων περιστατικών, έρευνες επισημαίνουν προβλήματα αστάθειας (instability), όπου μικρές αλλαγές στα δεδομένα οδηγούν σε διαφορετικές εξηγήσεις [50].

Οι Μηχανισμοί Προσοχής (Attention Mechanisms) στα μοντέλα NLP, όπως το BERT της παρούσας εργασίας, παρέχουν μια μορφή ερμηνευσιμότητας μέσω των βαρών προσοχής (Attention Weights). Αυτά τα βάρη δείχνουν σε ποιες λέξεις «εστίασε» το μοντέλο. Ωστόσο, μελέτες υποδεικνύουν ότι οι απλοί χάρτες προσοχής (attention maps) συχνά δεν ικανοποιούν τις ανάγκες της κλινικής πρακτικής, όπου προτιμώνται εξηγήσεις φυσικής γλώσσας ή ανάκτηση παρόμοιων περιστατικών [51].

Σημείωση: Οι τεχνικές SHAP και LIME παρουσιάζονται στο παρόν κεφάλαιο ως θεωρητικό πλαίσιο και δεν αποτελούν μέρος της τρέχουσας υλοποίησης. Η ενσωμάτωσή τους προτείνεται ως μελλοντική επέκταση (βλ. Ενότητα 7.4).

2.4.γ Υβριδικά Συστήματα και Διαφάνεια

Η έρευνα καταδεικνύει ότι τόσο η υπερβολική εμπιστοσύνη όσο και η δυσπιστία σε συστήματα AI μπορεί να είναι προβληματικές [53]. Για την αντιμετώπιση του ζητήματος της ευθύνης, σύγχρονες εφαρμογές όπως το Limbic Access και το Wysa υιοθετούν υβριδικές προσεγγίσεις [52]. Αντί να βασίζονται αποκλειστικά σε πιθανοτικά μοντέλα, ενσωματώνουν κανόνες ασφαλείας (rule-based guardrails) και παράγουν δομημένες αναφορές που επιτρέπουν την επαλήθευση [54].

Η παρούσα εργασία ακολουθεί αυτή την κατεύθυνση, υιοθετώντας μια Υβριδική Αρχιτεκτονική. Αντί να βασίζεται αποκλειστικά στη στοχαστική φύση του BERT, εισάγεται ένα ντετερμινιστικό επίπεδο κανόνων (Regex Safety Layer), επιδιώκοντας οι κρίσιμες ενδείξεις να είναι επαληθεύσιμες.

2.5 Προστασία Ιδιωτικότητας και GDPR στην Ψηφιακή Ψυχιατρική

Η επεξεργασία δεδομένων ψυχικής υγείας διέπεται από αυστηρούς ηθικούς και νομικούς κανόνες, με προεξάρχοντα τον Γενικό Κανονισμό για την Προστασία Δεδομένων (GDPR) [67]. Η ανάπτυξη αλγορίθμων ΤΝ υπόκειται στην αρχή Privacy by Design (Προστασία της Ιδιωτικότητας από τον Σχεδιασμό), όπως ορίζεται στο Άρθρο 25 του κανονισμού, επιβάλλοντας την ενσωμάτωση τεχνικών προστασίας ήδη από το στάδιο του σχεδιασμού.

2.5.α Κανονιστικό Πλαίσιο: Ευαίσθητα Δεδομένα και Δικαίωμα στην Εξήγηση

Τα δεδομένα ψυχικής υγείας εμπίπτουν στις «ειδικές κατηγορίες» του Άρθρου 9 GDPR, η επεξεργασία των οποίων απαιτεί αυξημένες εγγυήσεις. Παράλληλα, τα υποκείμενα των δεδομένων έχουν το «δικαίωμα στην εξήγηση» όσον αφορά την αυτοματοποιημένη λήψη αποφάσεων, σύμφωνα με τα Άρθρα 13-15. Αυτό δημιουργεί υποχρέωση για την αποφυγή χρήσης αδιαφανών συστημάτων χωρίς επεξήγηση, καθώς αυτό θα παραβίαζε την αρχή της διαφάνειας και της ενήμερης συγκατάθεσης [68].

2.5.β Τεχνικές Privacy-Preserving Machine Learning (PPML)

Για την εναρμόνιση της ανάγκης για δεδομένα με την προστασία της ιδιωτικότητας, η έρευνα έχει στραφεί σε τεχνικές Privacy-Preserving Machine Learning (PPML).

Η Ομοσπονδιακή Μάθηση (Federated Learning) επιτρέπει την εκπαίδευση μοντέλων απευθείας στις συσκευές των χρηστών, όπως κινητά τηλέφωνα, χωρίς τα δεδομένα να μεταφέρονται σε κεντρικούς διακομιστές. Αυτό είναι κατάλληλο για την ανάλυση ψηφιακών φαινοτύπων, όπως μοτίβα πληκτρολόγησης, χωρίς να εκτίθεται το περιεχόμενο των επικοινωνιών [66].

Η Διαφορική Ιδιωτικότητα (Differential Privacy) εισάγει στοχαστικό θόρυβο στα δεδομένα ή τις παραμέτρους του μοντέλου, διασφαλίζοντας μαθηματικά ότι η συνεισφορά ενός μεμονωμένου ατόμου δεν μπορεί να αναγνωριστεί. Συνδυάζεται συχνά με την Ομοσπονδιακή Μάθηση για την αποτροπή επιθέσεων αντιστροφής μοντέλου (model inversion) [66].

Η χρήση Συνθετικών Δεδομένων (Synthetic Data) μέσω αλγορίθμων, όπως GANs ή LLMs, επιτρέπει την παραγωγή τεχνητών δεδομένων που διατηρούν τις στατιστικές ιδιότητες των πραγματικών, χωρίς να αντιστοιχούν σε υπαρκτά πρόσωπα. Η προσέγγιση αυτή, την οποία υιοθετεί και η παρούσα εργασία μέσω των Gaussian Personas, επιχειρεί να αντιμετωπίσει το πρόβλημα της έλλειψης δεδομένων (data scarcity) και επιτρέπει την έρευνα χωρίς παραβίαση του GDPR [71].

2.5.γ Ηθικές Προκλήσεις και Εμπιστοσύνη Ασθενών

Η εφαρμογή της ΤΝ επηρεάζεται από την εμπιστοσύνη των χρηστών. Έρευνες δείχνουν ότι σημαντικό ποσοστό χρηστών ανησυχεί για λανθασμένες εκτιμήσεις από ΤΝ, ενώ η προθυμία μοιρασιάς πληροφοριών με chatbots είναι χαμηλότερη σε σύγκριση με ανθρώπους [75]. Παρουσιάζει ενδιαφέρον το εύρημα ότι χρήστες με σοβαρότερα προβλήματα τείνουν να είναι πιο διατεθειμένοι να δεχτούν ακριβή αλλά μη επεξηγήσιμη ΤΝ, γεγονός

που εγείρει ηθικά ζητήματα [75]. Συνεπώς, τα συστήματα TN θα πρέπει να σχεδιάζονται ως εργαλεία υποστήριξης που ενισχύουν τον επαγγελματία και όχι ως υποκατάστατα, ενσωματώνοντας διαφάνεια και λογοδοσία [70].

2.6 Η Χρήση Συνθετικών Δεδομένων στην Ιατρική Έρευνα

Η αυστηροποίηση του κανονιστικού πλαισίου για τα προσωπικά δεδομένα (GDPR) έχει δημιουργήσει εμπόδια στην εκπαίδευση μοντέλων AI στον ιατρικό τομέα [4]. Η έλλειψη ανοικτών συνόλων δεδομένων διαλόγου στα Ελληνικά είναι ιδιαίτερα αισθητή.

Η λύση που προκρίνεται στη σύγχρονη βιβλιογραφία είναι η χρήση Συνθετικών Δεδομένων (Synthetic Data). Μελέτες έχουν δείξει ότι μοντέλα που εκπαιδεύονται σε συνθετικά δεδομένα υψηλής ποιότητας, τα οποία παράγονται από Μεγάλα Γλωσσικά Μοντέλα (LLMs), μπορούν να γενικεύσουν σε πραγματικά δεδομένα, διατηρώντας παράλληλα την ιδιωτικότητα [5]. Ωστόσο, η διασφάλιση της εγκυρότητας αυτών των δεδομένων, ώστε να μην περιέχουν στατιστικές μεροληψίες (bias), παραμένει μια ανοικτή πρόκληση. Η παρούσα εργασία εισάγει τη μεθοδολογία των Gaussian Personas, συνδυάζοντας την πιθανοτική κατανομή με τη γλωσσική παραγωγή για τη δημιουργία προφίλ εκπαίδευσης [5].

2.7 Κριτική Αξιολόγηση και Κενά στη Βιβλιογραφία

Η συγκριτική ανάλυση των υφιστάμενων συστημάτων αναδεικνύει κενά που η παρούσα πτυχιακή εργασία επιχειρεί να καλύψει. Ο Πίνακας 2.1 παρουσιάζει μια ενδεικτική σύγκριση των χαρακτηριστικών των κυριότερων προσεγγίσεων, βασισμένη σε δημοσίως διαθέσιμη τεκμηρίωση. Θα πρέπει να σημειωθεί ότι πρόκειται για ποιοτική σύγκριση και όχι για εμπειρική αξιολόγηση απόδοσης.

Σύστημα	Γλώσσα	Μέθοδος	Επικύρωση	Περιορισμοί
Woebot	EN	Rule-based (CBT)	Κλινικές Δοκιμές	Περιορισμένη κατανόηση φυσικής γλώσσας, δεν υποστηρίζει Ελληνικά [14]
Wysa	EN	Hybrid AI	Αυτο-αναφορές	Έλλειψη σκορ βασισμένων σε τυποποιημένες κλίμακες (BDI/SCL), Συνδρομητικό μοντέλο [17]
Tess	EN	NLP/AI	Πιλοτικές μελέτες	Κλειστό σύστημα, Έλλειψη διαφάνειας αλγορίθμου [11]
Προτεινόμενο	EL	Hybrid (BERT+Regex)	Συνθετικά δεδομένα	Πρωτότυπο, απαιτεί επικύρωση σε πραγματικά δεδομένα

Πίνακας 2.1: Ενδεικτική Σύγκριση Συστημάτων (βασισμένη σε δημοσίως διαθέσιμη τεκμηρίωση)

Η προτεινόμενη υβριδική αρχιτεκτονική επιχειρεί να συνδυάσει την ευελιξία της μηχανικής μάθησης με την προβλεψιμότητα των κανόνων. Όπως αποτυπώνεται στο απόσπασμα κώδικα (Κώδικας 2.1), η προσέγγιση κανόνων (Regex) λειτουργεί ως ευρετικός μηχανισμός ελέγχου για τον εντοπισμό κρίσιμων ενδείξεων, ενώ η προσέγγιση BERT προσφέρει κατανόηση του συναισθηματικού περιεχομένου.

```
# --- ΜΕΡΟΣ Α: Προσέγγιση Κανόνων (Regex / Safety Layer) ---
# Στατική χαρτογράφηση ριζών για άμεση ανίχνευση κρίσιμων ενδείξεων
BECK_MAPPING = {
    r"\bαυτοκτον": "Ι. Αυτοκτονία", # Ανιχνεύει: αυτοκτονία, αυτοκτονήσω
    r"\bθάνατ": "Ι. Αυτοκτονία", # Ανιχνεύει: θάνατος, θανατηφόρο
    r"\bπεθάν": "Ι. Αυτοκτονία", # Ανιχνεύει: να πεθάνω
}

def detect_symptoms_advanced(text, mapping_dict):
    detected_items = set()

    # Κανονικοποίηση κειμένου (αφαίρεση τόνων, πεζά)
    text_normalized = strip_accents(text.lower())

    for pattern, description in mapping_dict.items():
        # Έλεγχος αν υπάρχει η ρίζα της λέξης-κλειδί (Hard Match)
        if re.search(pattern, text_normalized):

            # Έλεγχος πλαισίου: Υπάρχει άρνηση κοντά στη λέξη;
            # π.χ. Το "ΔΕΝ θέλω να πεθάνω" θα αγνοηθεί από το Regex Layer.
            if not check_negation_context(text_normalized, pattern, NEGATI
ONS):
                detected_items.add(description)

    return detected_items

# --- ΜΕΡΟΣ Β: Προσέγγιση Deep Learning (BERT) ---
# Στοχαστική πρόβλεψη βασισμένη σε συμφραζόμενα (Contextual Understanding)
def analyze_patient(tokenizer, model, text):

    # Μετατροπή κειμένου σε διανύσματα (Input IDs)
    inputs = tokenizer(text, return_tensors="pt", truncation=True, max_len
gth=512)

    with torch.no_grad():
        # Το μοντέλο παράγει Logits (πιθανότητες)
        logits = model(**inputs).logits

    # Εξαγωγή σκορ ακόμη και αν απουσιάζουν λέξεις-κλειδιά
    # π.χ. "Όλα μοιάζουν μάτια" -> Υψηλό σκορ κατάθλιψης
    predicted_score = logits[0].numpy()

    return predicted_score
```

Κώδικας 2.1: Απόσπασμα λογικής υβριδικής αρχιτεκτονικής (από Doctor.py)

Συμπερασματικά, τα κυριότερα ερευνητικά κενά που εντοπίζονται στη βιβλιογραφία αφορούν την έλλειψη συστημάτων που υποστηρίζουν την Ελληνική γλώσσα με κατανόηση πολιτισμικών συμφραζομένων, τις προκλήσεις στην ποιότητα των συνθετικών δεδομένων που συχνά δεν ελέγχονται επαρκώς ως προς τις στατιστικές τους ιδιότητες, καθώς και την απουσία μηχανισμών ελέγχου στα αμιγώς νευρωνικά συστήματα.

Κεφάλαιο 3: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ ΚΑΙ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ

3.1 Εισαγωγή και Ορισμοί

Η Τεχνητή Νοημοσύνη (TN) αποτελεί έναν εξελισσόμενο κλάδο της πληροφορικής που στοχεύει στη δημιουργία υπολογιστικών συστημάτων ικανών να προσομοιώνουν γνωστικές διαδικασίες, όπως η μάθηση, η εξαγωγή συμπερασμάτων και η αυτο-διόρθωση. Αν και οι ρίζες της θεωρίας εντοπίζονται στα μέσα του 20ου αιώνα με το έργο του Alan Turing σχετικά με τη νοημοσύνη των μηχανών [27], η σύγχρονη εφαρμογή της στην ιατρική εστιάζει κυρίως στην ικανότητα επεξεργασίας μεγάλου όγκου δεδομένων για την υποστήριξη αποφάσεων [10]. Στο πλαίσιο της παρούσας εργασίας, η TN δεν αντικαθιστά τον επαγγελματία ψυχικής υγείας, αλλά λειτουργεί επικουρικά, εντοπίζοντας μοτίβα στον ανθρώπινο λόγο που ενδέχεται να διαφεύγουν της προσοχής ή να απαιτούν χρονοβόρα ανάλυση.

3.2 Θεμελιώδεις Αρχές Νευρωνικών Δικτύων

3.2.α Ο Τεχνητός Νευρώνας και το Perceptron

Η λειτουργία των σύγχρονων αλγορίθμων Μηχανικής Μάθησης εμπνέεται από τον τρόπο μετάδοσης σημάτων στον ανθρώπινο εγκέφαλο. Η βασική δομική μονάδα είναι το Perceptron, το οποίο λειτουργεί ως ένας μαθηματικός προσομοιωτής ενός βιολογικού νευρώνα. Σε αντίθεση με τα βιολογικά δίκτυα, το Perceptron δέχεται αριθμητικές εισόδους, τις σταθμίζει μέσω βαρών (weights) και παράγει μια έξοδο εφόσον το άθροισμα υπερβεί ένα συγκεκριμένο κατώφλι ενεργοποίησης. Αυτή η διαδικασία, όταν κλιμακώνεται σε δίκτυα με πολλούς νευρώνες, επιτρέπει την επίλυση πολύπλοκων προβλημάτων ταξινόμησης. Η διαδικασία αυτή περιγράφεται μαθηματικά από την παρακάτω εξίσωση:

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right)$$

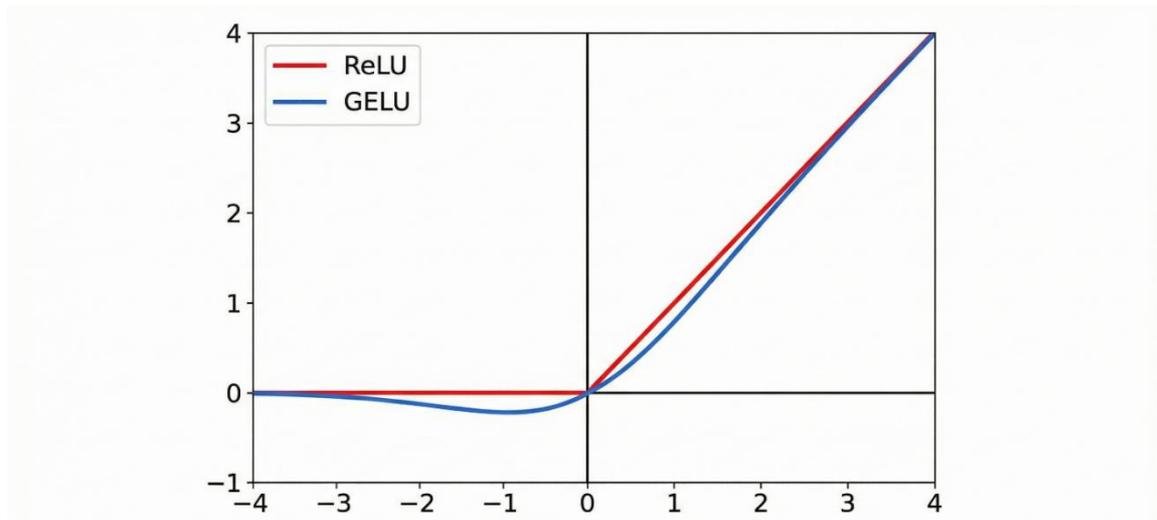
Όπου x_i αντιστοιχεί στις εισόδους του νευρώνα, w_i στα βάρη που καθορίζουν τη σημασία κάθε εισόδου, b στη σταθερά πόλωσης (bias), και f στη συνάρτηση ενεργοποίησης.

3.2.β Συναρτήσεις Ενεργοποίησης (Activation Functions)

Για να μπορέσει ένα νευρωνικό δίκτυο να μοντελοποιήσει μη γραμμικές σχέσεις στα δεδομένα, είναι απαραίτητη η χρήση μη γραμμικών συναρτήσεων ενεργοποίησης. Στην παρούσα υλοποίηση αξιοποιήθηκαν δύο κύριες συναρτήσεις.

Η συνάρτηση ReLU (Rectified Linear Unit) επιλέχθηκε για την υπολογιστική της αποδοτικότητα στα κρυφά επίπεδα του δικτύου, καθώς επιτρέπει ταχύτερη εκπαίδευση αποφεύγοντας το πρόβλημα της εξασθένισης της κλίσης [28]. Ορίζεται ως: $f(x) = \max(0, x)$.

Η συνάρτηση GELU (Gaussian Error Linear Unit) αποτελεί μια πιο πρόσφατη παραλλαγή που χρησιμοποιείται στην αρχιτεκτονική BERT για ομαλότερη ροή της πληροφορίας και καλύτερη πιθανοτική μοντελοποίηση. Όπως φαίνεται στην Εικόνα 3.1, η GELU προσφέρει ομαλότερη μετάβαση κοντά στο μηδέν σε σύγκριση με την ReLU.



Εικόνα 3.1: Σύγκριση Συναρτήσεων Ενεργοποίησης ReLU και GELU. Η GELU (μπλε) προσφέρει ομαλότερη ροή κλίσης σε σχέση με την απότομη ReLU (κόκκινο).

Σε επίπεδο κώδικα, η επιλογή της συνάρτησης ενεργοποίησης GELU γίνεται εγγενώς κατά την αρχικοποίηση του προ-εκπαιδευμένου μοντέλου. Όπως παρουσιάζεται στον Κώδικα 3.1, η φόρτωση του bert-base-greek-uncased-v1 μέσω της κλάσης AutoModelForSequenceClassification εισάγει αυτόματα τη ρύθμιση hidden_act="gelu" στα επίπεδα του Transformer, διασφαλίζοντας τη συμβατότητα με τα βάρη της προ-εκπαίδευσης.

```
from transformers import AutoModelForSequenceClassification

# Ορισμός του ονόματος του προ-εκπαιδευμένου μοντέλου.
# Το συγκεκριμένο μοντέλο (Greek BERT) χρησιμοποιεί αρχιτεκτονικά
# τη συνάρτηση GELU στα κρυφά επίπεδα (Encoder Layers).
model_name = "nlpraueb/bert-base-greek-uncased-v1"

# Φόρτωση της αρχιτεκτονικής για παλινδρόμηση (Regression)
# Η παράμετρος num_labels=1 προσαρμόζει την κεφαλή εξόδου για πρόβλεψη
# μίας τιμής.
model = AutoModelForSequenceClassification.from_pretrained(
    model_name,
    num_labels=1,
    problem_type="regression"
)
```

Κώδικας 3.1: Αρχικοποίηση μοντέλου με συνάρτηση GELU (από train_model_beck.py)

3.3 Επεξεργασία Φυσικής Γλώσσας: Από τα RNNs στους Transformers

3.3.α Η μετάβαση από τα RNNs στους Transformers

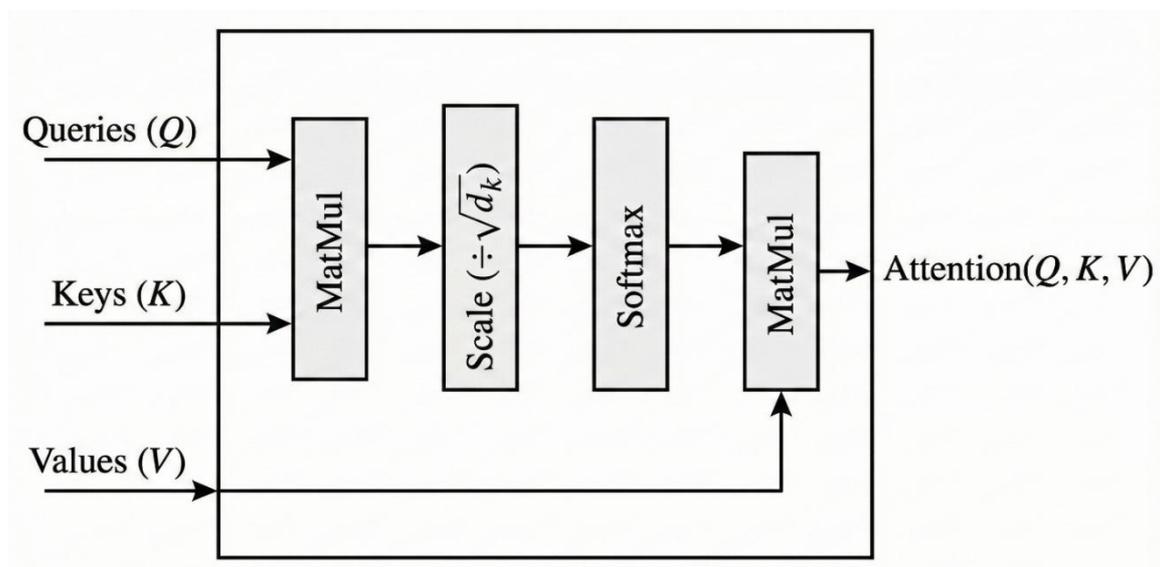
Μέχρι πρόσφατα, η ανάλυση κειμένου βασιζόταν κυρίως σε Αναδρομικά Νευρωνικά Δίκτυα (RNNs) και LSTMs. Ωστόσο, η εφαρμογή τους σε κείμενα παρουσίαζε σημαντικούς περιορισμούς [3]. Καθώς η επεξεργασία γινόταν σειριακά, το δίκτυο συχνά αδυνατούσε να διατηρήσει πληροφορίες από την αρχή της πρότασης όταν έφτανε στο τέλος, χάνοντας σημαντικές πληροφορίες. Επιπλέον, η αδυναμία παραλληλισμού της εκπαίδευσης καθιστούσε τη διαδικασία χρονοβόρα.

Η λύση δόθηκε με την εισαγωγή της αρχιτεκτονικής Transformer, η οποία επιτρέπει την ταυτόχρονη επεξεργασία όλης της πρότασης, εστιάζοντας επιλεκτικά στις λέξεις που έχουν τη μεγαλύτερη σημασία μέσω του μηχανισμού Attention [31].

3.3.β Μηχανισμός Αυτο-Προσοχής (Self-Attention)

Ο πυρήνας του Transformer είναι ο μηχανισμός Self-Attention, ο οποίος υπολογίζει τη σχετική σημασία κάθε λέξης σε σχέση με τις υπόλοιπες λέξεις της πρότασης. Η δομή του μηχανισμού παρουσιάζεται στην Εικόνα 3.2. Μαθηματικά, ορίζεται ως:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$



Εικόνα 3.2: Ο Μηχανισμός Scaled Dot-Product Attention

Στην εξίσωση αυτή, τα Q (Queries), K (Keys), V (Values) αποτελούν πίνακες που προκύπτουν από γραμμικούς μετασχηματισμούς της εισόδου, ενώ το d_k είναι η διάσταση των κλειδιών που χρησιμοποιείται για κλιμάκωση ώστε να αποφευχθούν πολύ μικρές κλίσεις.

Για παράδειγμα, στη φράση «Δεν νιώθω πια καμία χαρά», ο μηχανισμός συνδέει το αρνητικό μόριο «Δεν» με τη λέξη «χαρά». Αναγνωρίζει ότι η λέξη «χαρά» σε αυτό το πλαίσιο δεν έχει θετική έννοια, αλλά δηλώνει την απουσία της (ανηδονία), ένα σύμπτωμα της

κατάθλιψης. Αντίθετα, τα παλαιότερα δίκτυα RNNs, λόγω του προβλήματος της εξασθένισης της κλίσης, συχνά αδυνατούσαν να διατηρήσουν τη σχέση της άρνησης με τα επόμενα στοιχεία της πρότασης, οδηγώντας σε εσφαλμένη ταξινόμηση.

3.3.γ Πολυκεφαλικός Μηχανισμός Προσοχής (Multi-Head Attention)

Η χρήση μίας μόνο κεφαλής προσοχής περιορίζει την ικανότητα του μοντέλου να εστιάζει σε διαφορετικούς τύπους σχέσεων ταυτόχρονα, όπως συντακτικές έναντι σημασιολογικών. Το Multi-Head Attention επιλύει αυτό το πρόβλημα προβάλλοντας τα διανύσματα Q, K, V σε h διαφορετικούς υποχώρους αναπαράστασης [27]. Η εξίσωση ορίζεται ως:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O$$

όπου κάθε κεφαλή υπολογίζεται ανεξάρτητα: $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$.

Αυτή η αρχιτεκτονική επιτρέπει στο μοντέλο να επεξεργάζεται το κείμενο από διαφορετικές οπτικές ταυτόχρονα. Έρευνες στην ανάλυση μοντέλων BERT [55] έχουν δείξει ότι διαφορετικές κεφαλές εξειδικεύονται αυθόρμητα: ορισμένες εστιάζουν στη σύνταξη, όπως η σχέση ρήματος-αντικειμένου, ενώ άλλες σε σημασιολογικές αναφορές, όπως αντωνυμίες. Αυτό προσφέρει μια πολυδιάστατη κατανόηση του κειμένου που μπορεί να είναι χρήσιμη για τον εντοπισμό λεπτών γλωσσικών ενδείξεων.

3.3.δ Κωδικοποίηση Θέσης (Positional Encodings)

Επειδή η εξίσωση του Attention είναι αναλλοίωτη ως προς τη μετάθεση (permutation invariant) και οι Transformers δεν διαθέτουν αναδρομικότητα, η πληροφορία της σειράς των λέξεων πρέπει να εισαχθεί ρητά. Αυτό επιτυγχάνεται μέσω των Positional Encodings, χρησιμοποιώντας ημιτονοειδείς συναρτήσεις διαφορετικών συχνοτήτων [27]:

$$PE_{(pos, 2i)} = \sin(pos/10000^{2i/d_{model}})$$

$$PE_{(pos, 2i+1)} = \cos(pos/10000^{2i/d_{model}})$$

Η επιλογή αυτή επιτρέπει στο μοντέλο να μάθει τις σχετικές αποστάσεις μεταξύ των λέξεων, καθώς για κάθε σταθερή μετατόπιση k , το PE_{pos+k} μπορεί να εκφραστεί ως γραμμικός μετασχηματισμός του PE_{pos} . Αυτό είναι σημαντικό για την ανίχνευση της εμβέλειας της άρνησης (negation scope), διακρίνοντας για παράδειγμα το «δεν θέλω να πεθάνω» από το «θέλω να πεθάνω» [56].

3.3.ε Εφαρμογές στην Υπολογιστική Ψυχιατρική

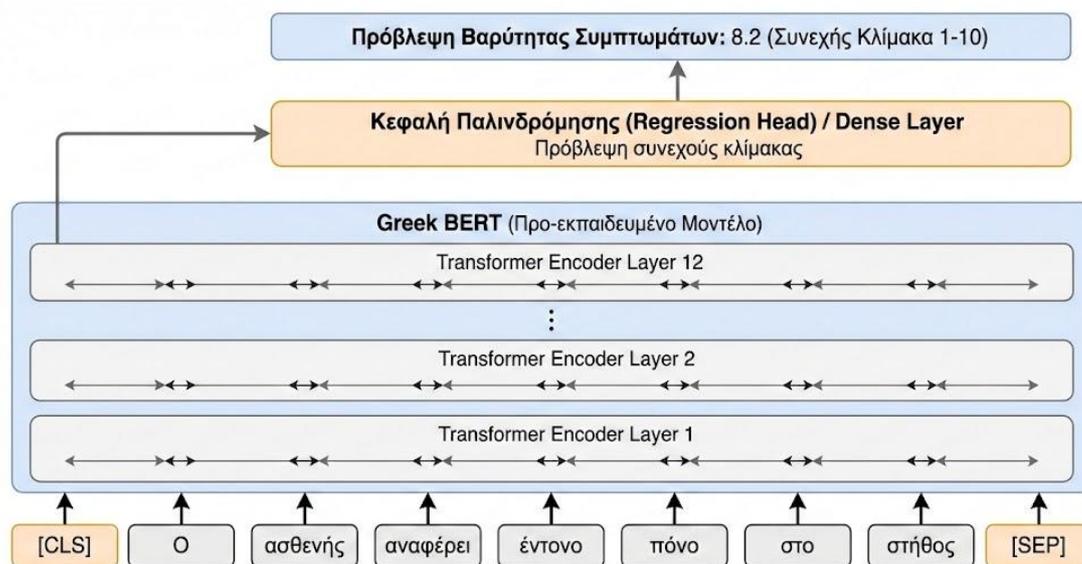
Η εφαρμογή του Multi-Head Attention στην ανάλυση κειμένων σχετικών με την ψυχική υγεία έχει οδηγήσει στην ανάπτυξη εξειδικευμένων μοντέλων. Το RAMHA [57] συνδυάζει έναν κωδικοποιητή RoBERTa με μηχανισμούς adapters, επιδεικνύοντας καλή απόδοση στην ταξινόμηση συναισθημάτων μέσω της εστίασης τόσο σε στατικά όσο και σε χρονικά χαρακτηριστικά του κειμένου. Το AttentionDep [58] ενσωματώνει εξωτερική

γνώση από γράφους γνώσης στη διαδικασία του attention, βελτιώνοντας την ευαισθησία σε συμπτώματα σοβαρότητας. Επιπλέον, εξειδικευμένες κεφαλές προσοχής έχει αποδειχθεί ότι παίζουν σημαντικό ρόλο στη σύνδεση αρνητικών μορίων με τα στοιχεία που τροποποιούν [56], μειώνοντας τα ψευδώς θετικά αποτελέσματα κατά την ανάλυση κειμένων.

3.4 Μοντέλα BERT και Προσαρμογή

Το μοντέλο BERT (Bidirectional Encoder Representations from Transformers) διαφοροποιείται από προηγούμενες προσεγγίσεις καθώς αναλύει το κείμενο αμφίδρομα, λαμβάνοντας υπόψη τόσο το προηγούμενο όσο και το επόμενο περιβάλλον κάθε λέξης. Στην παρούσα εργασία, αξιοποιήθηκε το Greek BERT [18], το οποίο έχει προ-εκπαιδευτεί σε μεγάλα σώματα ελληνικών κειμένων. Η προσέγγιση που ακολουθήθηκε βασίζεται στην τεχνική Fine-Tuning for Regression: αντί για απλή κατηγοριοποίηση, όπως Θετικό/Αρνητικό, προστέθηκε ένα τελικό επίπεδο που εκπαιδεύεται να προβλέπει τη βαρύτητα των συμπτωμάτων σε συνεχή αριθμητική κλίμακα.

Στην Εικόνα 3.3 απεικονίζεται η αρχιτεκτονική του συστήματος, όπου διακρίνεται η ροή από το Greek BERT στην κεφαλή παλινδρόμησης.



Εικόνα 3.3: Η αρχιτεκτονική του Greek BERT προσαρμοσμένη για παλινδρόμηση (Regression Fine-Tuning)

Αυτή η προσέγγιση επιτρέπει στο σύστημα να προβλέπει τη βαρύτητα ενός συμπτώματος σε μια συνεχή κλίμακα, όπως 0 έως 4, προσφέροντας μια λεπτομερή εικόνα αντί για μια απλή δυαδική ταξινόμηση.

3.4.α Διαδικασία Fine-Tuning

Το προ-εκπαιδευμένο Greek BERT προσαρμόστηκε μέσω των ακόλουθων βημάτων.

Πρώτον, πραγματοποιήθηκε αφαίρεση του Classification Head. Το τελικό επίπεδο του BERT που προβλέπει γενικές κατηγορίες, όπως Next Sentence Prediction, αφαιρέθηκε και αντικαταστάθηκε με ένα απλό γραμμικό επίπεδο (linear layer) που εξάγει συνεχείς τιμές.

Δεύτερον, εφαρμόστηκε κλειδωμα των πρώτων επιπέδων (Freezing). Τα πρώτα 8 από τα 12 επίπεδα του Encoder παρέμειναν ανενεργά (frozen), διατηρώντας τη γενική γλωσσική γνώση όπως λεξιλόγιο και γραμματική. Θα πρέπει να σημειωθεί ότι η επιλογή του αριθμού των κλειδωμένων επιπέδων (8 από 12) αποτελεί σχεδιαστική απόφαση που βασίστηκε σε πρακτικές δοκιμές και δεν υποστηρίζεται από συστηματική μελέτη αφαίρεσης (ablation study). Μελλοντική εργασία θα μπορούσε να διερευνήσει την επίδραση διαφορετικών διαμορφώσεων.

Τρίτον, πραγματοποιήθηκε εκπαίδευση των τελικών επιπέδων. Μόνο τα τελευταία 4 επίπεδα και το νέο regression head εκπαιδεύτηκαν (fine-tuned) στα παραγόμενα συνθετικά δεδομένα.

Αυτή η στρατηγική επιλέχθηκε για να αποτρέψει το φαινόμενο της «καταστροφικής λήθης» (catastrophic forgetting), όπου το μοντέλο, στην προσπάθειά του να μάθει τα νέα δεδομένα, θα χάνει τις θεμελιώδεις δομές της Ελληνικής γλώσσας.

3.5 Βελτιστοποίηση και Αξιολόγηση: AdamW και Κλινικές Μετρικές

Η εκπαίδευση μοντέλων Transformers εξαρτάται από δύο παράγοντες: τον αλγόριθμο βελτιστοποίησης και τις μετρικές αξιολόγησης. Η επιλογή αυτών των εργαλείων καθορίζει όχι μόνο την τεχνική ακρίβεια, αλλά και την αξιοπιστία του συστήματος.

3.5.α Οπισθοδιάδοση (Backpropagation)

Πριν την εφαρμογή του βελτιστοποιητή, προηγείται ο υπολογισμός των κλίσεων. Η διαδικασία ακολουθεί δύο επαναλαμβανόμενες φάσεις. Κατά το Forward Pass, το κείμενο διέρχεται από το δίκτυο και παράγεται μια πρόβλεψη. Κατά το Backward Pass, το σφάλμα διαδίδεται προς τα πίσω, ενημερώνοντας τα βάρη ανάλογα με τη συνεισφορά τους στο τελικό λάθος.

3.5.β Ο Αλγόριθμος AdamW

Για πολλά χρόνια, ο αλγόριθμος Adam (Adaptive Moment Estimation) χρησιμοποιούνταν ευρέως λόγω της ταχύτητάς του. Ωστόσο, παρατηρήθηκε ότι συχνά οδηγούσε σε λύσεις που γενίκευαν χειρότερα από τον απλό SGD, ειδικά σε βαθιά δίκτυα [72]. Το πρόβλημα εντοπίστηκε στον τρόπο εφαρμογής της κανονικοποίησης L2 (L2 regularization) εντός του προσαρμοστικού μηχανισμού του Adam [71].

Ο αλγόριθμος AdamW αντιμετώπισε αυτό το πρόβλημα μέσω της αποσύζευξης (decoupling) της μείωσης βαρών (weight decay). Αντί να προσθέτει την ποινή L2 στην κλίση, το AdamW την εφαρμόζει απευθείας στις παραμέτρους, ανεξάρτητα από τον υπολογισμό των ροπών [72]. Αυτή η αλλαγή είναι σημαντική για τους Transformers, καθώς επιτρέπει στο μοντέλο να συνεχίσει να μαθαίνει αποτελεσματικά ακόμα και όταν τα βάρη είναι μεγάλα, αποτρέποντας την πρόωρη στασιμότητα και επιτυγχάνοντας καλύτερη γενίκευση [72].

3.5.γ Κλινική Αξιολόγηση: MSE/RMSE vs MAE/R²

Στην πρόβλεψη της σοβαρότητας συμπτωμάτων, όπως το σκορ BDI-II (0-63), η επιλογή της συνάρτησης σφάλματος έχει πρακτικές συνέπειες [76].

Το RMSE (Root Mean Squared Error) επιβαρύνει δυσανάλογα τα μεγάλα σφάλματα. Αυτό είναι επιθυμητό όταν θέλουμε να αποφύγουμε σημαντικές αστοχίες, όπως πρόβλεψη χαμηλής βαρύτητας σε περίπτωση υψηλής βαρύτητας. Στη βιβλιογραφία, ένα RMSE < 5.0 θεωρείται καλό επίπεδο απόδοσης για την κλίμακα BDI-II, καθώς αντιστοιχεί στην Ελάχιστη Κλινικά Σημαντική Διαφορά (MCID) [79].

Το MAE (Mean Absolute Error) αντιμετωπίζει όλα τα σφάλματα ισότιμα. Ενώ είναι πιο ανθεκτικό στα outliers, μπορεί να υποτιμήσει τον κίνδυνο σε κρίσιμες περιπτώσεις .

Το Brier Score χρησιμοποιείται για την πρόβλεψη πιθανοτήτων και αξιολογεί τη βαθμονόμηση (calibration), διασφαλίζοντας ότι η προβλεπόμενη πιθανότητα αντιστοιχεί στην πραγματική συχνότητα του συμβάντος .

Η παρούσα εργασία υιοθετεί τον AdamW για τη βελτιστοποίηση λόγω της καταλληλότητάς του για αρχιτεκτονικές Transformer και το RMSE ως κύριο μετρικό αξιολόγησης.

3.6 Διαδικασία Εκπαίδευσης και Βελτιστοποίηση

Η εκπαίδευση του μοντέλου πραγματοποιήθηκε με πρωτόκολλα για την αποφυγή της υπερ-προσαρμογής (overfitting), δεδομένου του μεγέθους των δεδομένων.

3.6.α Συναρτήσεις Κόστους

Για την εκπαίδευση του μοντέλου παλινδρόμησης, επιλέχθηκε το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error - MSE), το οποίο υπολογίζει το μέσο όρο των τετραγωνικών διαφορών μεταξύ των προβλεπόμενων και των πραγματικών τιμών:

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

Η τετραγωνική φύση της συνάρτησης επιβαρύνει περισσότερο τα μεγάλα σφάλματα σε σχέση με τα μικρά. Αυτό καθιστά το μοντέλο ευαίσθητο σε ακραίες αποκλίσεις (outliers), κάτι που μπορεί να είναι επιθυμητό στον εντοπισμό σημαντικών αποκλίσεων, όπου η αποτυχία εντοπισμού σοβαρών συμπτωμάτων έχει μεγαλύτερο κόστος από την εσφαλμένη εκτίμηση ήπιων περιπτώσεων.

3.6.β Οπισθοδιάδοση (Backpropagation)

Η βελτιστοποίηση των βαρών του δικτύου επιτυγχάνεται μέσω του αλγορίθμου οπισθοδιάδοσης, ο οποίος υπολογίζει τις κλίσεις (gradients) της συνάρτησης κόστους ως προς κάθε βάρους. Η διαδικασία ακολουθεί δύο φάσεις. Κατά το Forward Pass, το κείμενο διέρχεται από το δίκτυο επίπεδο-προς-επίπεδο και παράγεται μια πρόβλεψη. Κατά το Backward Pass, το σφάλμα διαδίδεται προς τα πίσω, από την έξοδο προς την είσοδο, ενημερώνοντας τα βάρη ανάλογα με τη συνεισφορά τους στο τελικό λάθος.

3.7.β Επαύξηση Δεδομένων (Data Augmentation)

Για την αντιμετώπιση της έλλειψης μεγάλου όγκου δεδομένων και της ανισορροπίας που αναφέρθηκε, αναπτύχθηκε μηχανισμός Data Augmentation. Στο Παράθεμα 3.2 φαίνεται η κλάση AugmentedClinicalDataset, η οποία εφαρμόζει τυχαίες αλλοιώσεις (π.χ. αφαίρεση τόνων) σε πραγματικό χρόνο.

```
def augment_greek_text(text, intensity=0.3):
    """
    Εφαρμόζει τυχαίες αλλοιώσεις στο ελληνικό κείμενο για robust training.
    Π.χ. "άγχος" -> "αγχος" (αφαίρεση τόνου) ή "είμαι" -> "ειμαι"
    """
    if random.random() < intensity:
        # Αφαίρεση τόνου από τυχαίο χαρακτήρα (κοινό λάθος σε mobile typin
g)
        text = remove_random_accent(text)
    return text

class AugmentedClinicalDataset(torch.utils.data.Dataset):
    """
    Προσαρμοσμένη κλάση PyTorch Dataset.
    Ενσωματώνει τη διαδικασία Tokenization και Augmentation στη ροή εκπαίδ
ευσης.
    """
    def __init__(self, texts, labels, tokenizer, augment=False):
        self.texts = texts
        self.labels = labels
        self.tokenizer = tokenizer
        self.augment = augment # Flag ενεργοποίησης augmentation

    def __getitem__(self, idx):
        text = str(self.texts[idx])

        # Εφαρμογή Augmentation ΜΟΝΟ στο σύνολο εκπαίδευσης
        if self.augment:
            text = augment_greek_text(text, intensity=0.3)

        # Tokenization για το BERT (padding/truncation σε 512 tokens)
        encoding = self.tokenizer(
            text,
            truncation=True,
            padding='max_length',
            max_length=512,
            return_tensors='pt'
        )

        # Προετοιμασία Tensors για την GPU
        item = {key: val.squeeze(0) for key, val in encoding.items()}
        item['labels'] = torch.tensor(self.labels[idx], dtype=torch.float)

    return item
```

Κώδικας 3.2: Υλοποίηση Dataset με On-the-fly Greek Data Augmentation (από το αρχείο *data_toolkt.py* και *train_model_beck.py*)

3.7.γ Αλγόριθμος AdamW και Κανονικοποίηση

Για τη βελτιστοποίηση των βαρών επιλέχθηκε ο αλγόριθμος AdamW. Πρόκειται για μια παραλλαγή του Adam που διαχωρίζει τη μείωση βάρους (weight decay) από τον υπολογισμό της κλίσης:

$$\theta_{t+1} = \theta_t - \eta \left(\frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} + \lambda \theta_t \right)$$

Όπου λ είναι ο συντελεστής weight decay (0.01) και ϵ ο ρυθμός μάθησης. Παράλληλα, για την ενίσχυση της ικανότητας γενίκευσης του μοντέλου, εφαρμόστηκαν τεχνικές κανονικοποίησης (Regularization). Συγκεκριμένα, χρησιμοποιήθηκε η τεχνική Dropout με πιθανότητα $p = 0.1$ για την απενεργοποίηση τυχαίων νευρώνων, καθώς και η τεχνική Early Stopping για τη διακοπή της εκπαίδευσης όταν το σφάλμα στο σύνολο επαλήθευσης σταματήσει να μειώνεται.

3.7.δ Παραμετροποίηση και Παρακολούθηση Σύγκλισης

Η διαδικασία εκπαίδευσης υλοποιήθηκε μέσω της βιβλιοθήκης Hugging Face Transformers. Στον Κώδικα 3.3 παρουσιάζεται η διαμόρφωση των παραμέτρων εκπαίδευσης (TrainingArguments), όπου ορίζονται το πλήθος των εποχών, το μέγεθος δέσμης (batch size) και η στρατηγική αξιολόγησης.

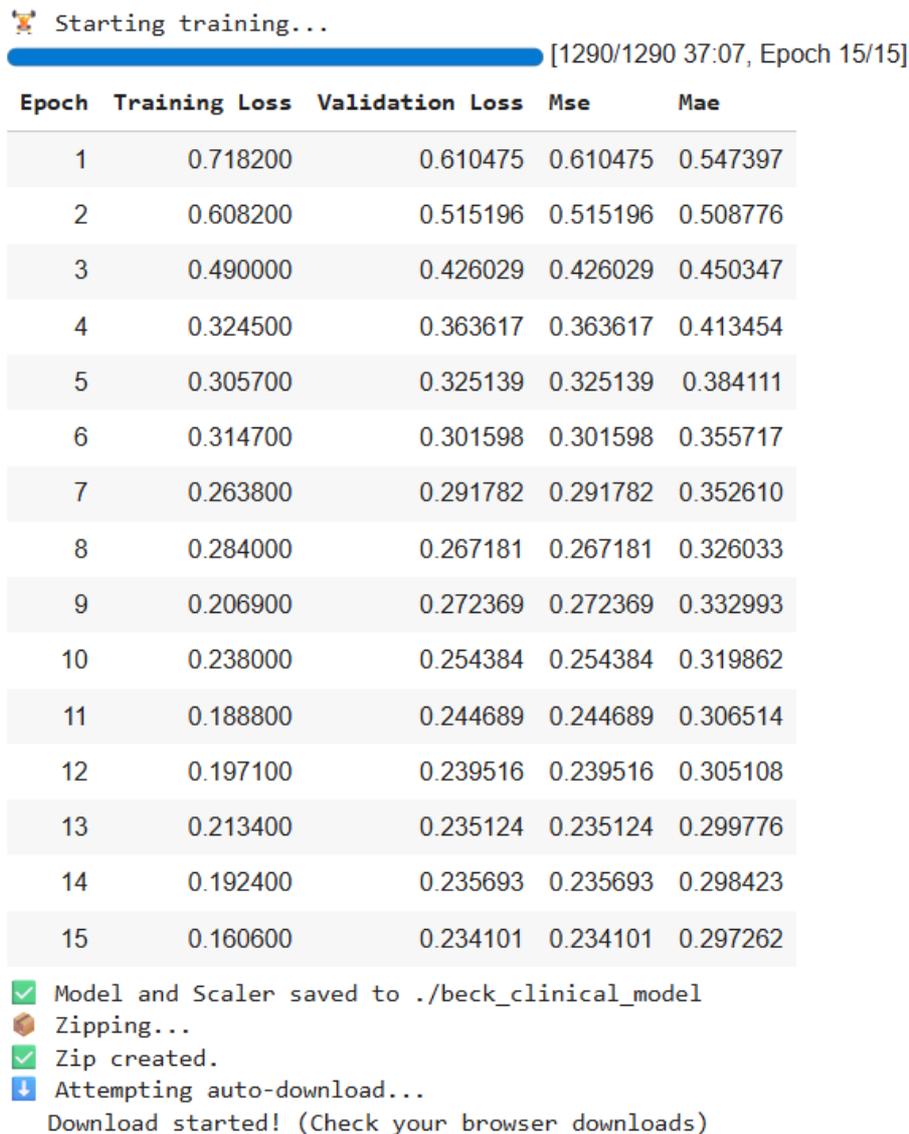
Για τη βέλτιστη διαχείριση των υπολογιστικών πόρων και του χώρου αποθήκευσης, ενεργοποιήθηκε η στρατηγική «κυλιόμενων σημείων ελέγχου» (rolling checkpoints). Μέσω της παραμέτρου `save_total_limit=2`, το σύστημα διατηρεί αποθηκευμένα στον δίσκο μόνο τα δύο πιο πρόσφατα μοντέλα. Επιπλέον, σημαντική για την αποφυγή της υπερπροσαρμογής είναι η ενεργοποίηση της παραμέτρου `load_best_model_at_end=True`. Με αυτή τη ρύθμιση, το σύστημα στο τέλος της διαδικασίας απορρίπτει την τελευταία κατάσταση των βαρών και επαναφέρει αυτόματα την έκδοση του μοντέλου που πέτυχε το ελάχιστο Μέσο Τετραγωνικό Σφάλμα (MSE) στο σύνολο επαλήθευσης.

```
# Ορισμός υπερ-παραμέτρων εκπαίδευσης (Training Arguments)
training_args = TrainingArguments(
    output_dir='./results_beck_items',
    num_train_epochs=15,           # Μέγιστος αριθμός εποχών
    per_device_train_batch_size=16, # Μέγεθος δέσμης εκπαίδευσης
    per_device_eval_batch_size=16,  # Μέγεθος δέσμης αξιολόγησης
    warmup_steps=100,             # Βήματα προθέρμανσης (Learning Rate
    # Warmup)
    weight_decay=0.01,           # Συντελεστής μείωσης βάρους (Regula
    # rization)
    learning_rate=2e-5,          # Ρυθμός μάθησης
    logging_steps=10,
    eval_strategy="epoch",       # Αξιολόγηση στο τέλος κάθε εποχής
    save_strategy="epoch",       # Αποθήκευση στο τέλος κάθε εποχής
    load_best_model_at_end=True,  # Φόρτωση του βέλτιστου μοντέλου στο
    # τέλος
    metric_for_best_model="mse",  # Κριτήριο βελτιστοποίησης (Mean Squ
    # ared Error)
    greater_is_better=False      # Στόχος η ελαχιστοποίηση του σφάλμα
```

τος
)

Κώδικας 3.3: Παράμετροι εκπαίδευσης και βελτιστοποίησης (από `train_model_beck.py`).

Τα αποτελέσματα της εκπαίδευσης απεικονίζονται στην Εικόνα 3.5. Παρατηρείται ότι τόσο το σφάλμα εκπαίδευσης (Training Loss) όσο και το σφάλμα επικύρωσης (Validation Loss) μειώνονται σταθερά, ενώ η μεταξύ τους απόκλιση παραμένει μικρή, γεγονός που αποτελεί ένδειξη επιτυχούς αποφυγής του φαινομένου της υπερ-προσαρμογής.



Εικόνα 3.5: Καμπύλες μάθησης (*Training vs Validation Loss*)

3.8 Πλαίσια Ανάπτυξης και Βιβλιοθήκες (Frameworks & Libraries)

Για την υλοποίηση του συστήματος χρησιμοποιήθηκε ένα σύνολο σύγχρονων εργαλείων ανοικτού κώδικα, καλύπτοντας τρία επίπεδα: την εκπαίδευση του μοντέλου (AI), την εξυπηρέτηση των αιτημάτων (Backend) και τη διεπαφή χρήστη (Mobile Frontend). Η πλήρης υλοποίηση του συστήματος, συμπεριλαμβανομένων των scripts εκπαίδευσης και της εφαρμογής Android, είναι ελεύθερα προσβάσιμη στο αποθετήριο της εργασίας [88].

Το PyTorch επιλέχθηκε ως το βασικό πλαίσιο βαθιάς μάθησης λόγω της δυναμικής του φύσης (dynamic computational graph) και της ευρείας αποδοχής του στην ερευνητική κοινότητα [76]. Χρησιμοποιήθηκε για τη φόρτωση των τανυστών και την εκτέλεση του αλγορίθμου οπισθοδιάδοσης. Σε συνδυασμό με αυτό, η βιβλιοθήκη Hugging Face Transformers αξιοποιήθηκε για την άντληση του προ-εκπαιδευμένου μοντέλου bert-base-greek-uncased-v1 και των αντιστοιχών Tokenizers, απλοποιώντας σημαντικά τη διαδικασία Fine-Tuning.

Για την ανάπτυξη του εξυπηρετητή (Backend) επιλέχθηκε το FastAPI αντί του παραδοσιακού Flask. Πρόκειται για ένα σύγχρονο web framework που υποστηρίζει ασύγχρονη εκτέλεση (async/await), επιτρέποντας την ταυτόχρονη εξυπηρέτηση πολλαπλών χρηστών χωρίς καθυστερήσεις στην απόκριση του μοντέλου AI. Η εφαρμογή πελάτη (Client) αναπτύχθηκε για λειτουργικό σύστημα Android με χρήση της γλώσσας Kotlin και του εργαλείου Jetpack Compose, το οποίο επιτρέπει δηλωτική (declarative) κατασκευή του UI. Τέλος, η βιβλιοθήκη OpenPyXL της Python χρησιμοποιήθηκε για την αυτοματοποιημένη παραγωγή αναλυτικών αναφορών σε μορφή Excel, επιτρέποντας την άμεση αξιοποίηση των αποτελεσμάτων από τους επαγγελματίες ψυχικής υγείας.

Στον Κώδικα 3.4 παρατίθενται οι ακριβείς εκδόσεις των βιβλιοθηκών που χρησιμοποιήθηκαν στο περιβάλλον Python, διασφαλίζοντας την επαναληψιμότητα (reproducibility) της πειραματικής διαδικασίας.

```
# Web Server & API
fastapi==0.109.0
uvicorn[standard]==0.27.0
slowapi==0.1.9 # Rate Limiting για προστασία του server

# Machine Learning & Data Processing
torch==2.1.2
transformers==4.36.2
scikit-learn==1.4.0 # Metrics & Splitting
numpy==1.26.3
openpyxl==3.1.2 # Excel Report Generation
```

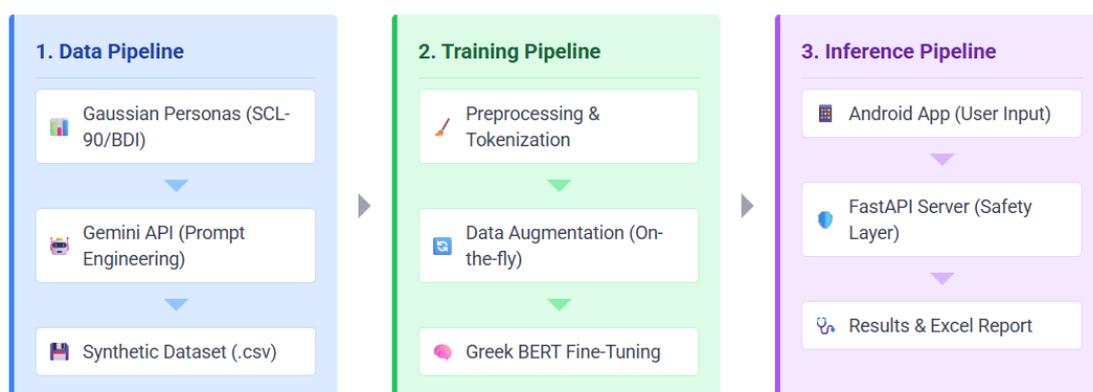
Κώδικας 3.4: Το αρχείο εξαρτήσεων requirements.txt του συστήματος.

Κεφάλαιο 4: ΜΕΘΟΔΟΛΟΓΙΑ ΥΛΟΠΟΙΗΣΗΣ ΚΑΙ ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΣΥΣΤΗΜΑΤΟΣ

4.1 Εισαγωγή στη Ροή Εργασίας (System Pipeline)

Η ανάπτυξη ενός συστήματος mHealth που βασίζεται σε Τεχνητή Νοημοσύνη απαιτεί τον συντονισμό πολλαπλών υποσυστημάτων. Όπως οι διαδικασίες αναγνώρισης προτύπων απαιτούν προεπεξεργασία δεδομένων, έτσι και η παρούσα προσέγγιση απαιτεί μια δομημένη ροή εργασίας για τη διαχείριση κειμένου και την εξαγωγή εκτιμήσεων.

Η ροή εργασίας (Pipeline) του συστήματος αποτελείται από τρία κύρια στάδια, όπως απεικονίζεται στην Εικόνα 4.1. Το πρώτο στάδιο αφορά την Παραγωγή και Προετοιμασία Δεδομένων (Data Pipeline), όπου δημιουργούνται συνθετικά σενάρια για σκοπούς εκπαίδευσης. Το δεύτερο στάδιο περιλαμβάνει την Εκπαίδευση Μοντέλων (Training Pipeline), όπου πραγματοποιείται το Fine-tuning του Greek BERT με τεχνικές Augmentation. Το τρίτο στάδιο αφορά την Εκτέλεση και τον Συμπερασμό (Inference Pipeline), δηλαδή τη λειτουργία του Server και της Android εφαρμογής σε πραγματικό χρόνο.



Εικόνα 4.1: Η αρχιτεκτονική ροή του συστήματος από την παραγωγή δεδομένων έως την τελική εκτίμηση.

Η υβριδική μηχανή συμπερασμού αποτελεί προσαρμογή του μοντέλου του Βιολέντη [30] για ανίχνευση δυσλεξίας. Στην αρχική μελέτη συνδυάστηκαν CNNs με ευρετικούς αλγορίθμους κατάτμησης εικόνας. Στην παρούσα εργασία, η ίδια αρχή μεταφέρεται στην Επεξεργασία Φυσικής Γλώσσας: το οπτικό επίπεδο αντικαταστάθηκε από το μοντέλο BERT, ενώ οι ευρετικοί κανόνες αντικαταστάθηκαν από ένα επίπεδο Regex για εντοπισμό ενδείξεων κινδύνου. Η προσέγγιση διατηρεί την αρχή της διπλής επαλήθευσης, προσαρμοσμένη στα ψυχομετρικά δεδομένα.

4.2 Μεθοδολογία Παραγωγής Συνθετικών Δεδομένων

Λόγω της έλλειψης πραγματικών δεδομένων εξαιτίας του GDPR, σχεδιάστηκε μηχανισμός παραγωγής συνθετικών δεδομένων. Τα δεδομένα αυτά αποτελούν τεχνική απόδειξη σκοπιμότητας (proof-of-concept) και δεν έχουν επικυρωθεί ως προς την οικολογική τους εγκυρότητα με σύγκριση με πραγματικά κλινικά δεδομένα.

4.2.α Στοχαστική Μοντελοποίηση Κλινικών Προφίλ (Gaussian Personas)

Η μέθοδος βασίζεται στην υπόθεση ότι η τάση ενός χρήστη να επιλέξει μια βαθμίδα στην κλίμακα Likert (0-4) δεν είναι τυχαία, αλλά ακολουθεί κατανομή πιθανοτήτων γύρω από μια μέση τιμή που αντιστοιχεί στην κατάσταση του.

Προσαρμογή στην Κατανομή Φοιτητικού Πληθυσμού

Η βιβλιογραφία δείχνει ότι τα σκορ BDI-II σε μη κλινικούς φοιτητικούς πληθυσμούς παρουσιάζουν θετική λοξότητα, με την πλειοψηφία να συγκεντρώνεται στις χαμηλές τιμές (0-13) [40]. Στην παρούσα εργασία επιλέχθηκε η Κανονική Κατανομή ως απλοποιημένη προσέγγιση. Για να αντισταθμιστεί η διαφορά, οι παράμετροι (μ , σ) ρυθμίστηκαν ώστε να προσομοιώνουν συγκέντρωση προς τις χαμηλές τιμές για την κατηγορία «Minimal». Αυτή η απλοποίηση αποτελεί περιορισμό της μεθοδολογίας.

Βαρύτητα και Συχνότητα Συμπτωμάτων

Η παραγωγή των συνθετικών δεδομένων έλαβε υπόψη τα ευρήματα της μελέτης των de Sá Junior et al. (2019), τα οποία καταδεικνύουν ότι δεν έχουν όλα τα συμπτώματα την ίδια συχνότητα εμφάνισης στους φοιτητές [42]. Τα σωματικά συμπτώματα, όπως οι «Αλλαγές στον Ύπνο» (Item 16) και η «Κόπωση» (Item 20), εμφανίζουν υψηλά ποσοστά αναφοράς, συχνά λόγω του ακαδημαϊκού φόρτου και όχι απαραίτητα λόγω κατάθλιψης [42]. Αντίθετα, συμπτώματα όπως ο «Αυτοκτονικός Ιδεασμός» (Item 9) είναι σπάνια στο γενικό φοιτητικό πληθυσμό (μέσο σκορ περίπου 0.08), εμφανιζόμενα κυρίως σε περιπτώσεις σοβαρής διαταραχής [42]. Αντίστοιχα, για την κλίμακα SCL-90-R, λαμβάνεται υπόψη ότι οι υψηλές τιμές στην υποκλίμακα Obsessive-Compulsive σε φοιτητικά δείγματα συχνά αντικατοπτρίζουν γνωστικό φόρτο και ακαδημαϊκό άγχος, παρά κλινική ιδεοψυχαναγκαστική διαταραχή [44].

Αυτή η ανομοιογένεια στη συχνότητα των συμπτωμάτων (class imbalance) αποτελεί έναν από τους λόγους για τη χρήση της Υβριδικής Αρχιτεκτονικής που περιγράφεται στην Ενότητα 5.2.2: το νευρωνικό δίκτυο (BERT) εκπαιδεύεται να αναγνωρίζει τα συχνά, ασαφή συμπτώματα (όπως η κόπωση), ενώ το επίπεδο ελέγχου (Regex Safety Layer) στοχεύει στον εντοπισμό των σπάνιων αλλά σημαντικών ενδείξεων (όπως η αυτοκτονία), οι οποίες ενδέχεται να υπο-εκπροσωπούνται στα δεδομένα εκπαίδευσης.

Η πιθανότητα P ένας προσομοιωμένος χρήστης της κατηγορίας C να επιλέξει την απάντηση $k \in \{0,1,2,3,4\}$ υπολογίζεται μέσω της ολοκλήρωσης της κανονικής κατανομής:

$$P(X = k|C) = \frac{1}{Z} \int_{k-0.5}^{k+0.5} \frac{1}{\sigma_c \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu_c}{\sigma_c}\right)^2} dx$$

Στην εξίσωση αυτή, το μ_c (Μέση Τιμή) καθορίζει το κέντρο βάρους των απαντήσεων του χρήστη. Για παράδειγμα, στην κατηγορία Severe, το μ ορίζεται στο 3.5, ωθώντας τις πιθανότητες προς τις επιλογές «3» και «4». Το σ_c (Τυπική Απόκλιση) ελέγχει τη μεταβλητότητα των απαντήσεων, όπου μια μικρή τιμή σ υποδηλώνει συνεπή συμπεριφορά, ενώ μια μεγαλύτερη επιτρέπει την εμφάνιση θορύβου (outliers), προσομοιώνοντας την

ανθρώπινη ασυνέπεια. Το Z (Παράγοντας Κανονικοποίησης) διασφαλίζει ότι το άθροισμα των πιθανοτήτων για τις 5 επιλογές ισούται με 1.

Η υλοποίηση του αλγορίθμου παρουσιάζεται στον Κώδικα 4.1.

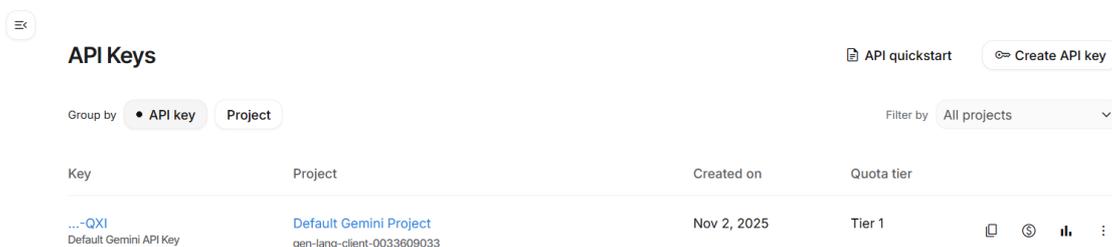
```
def get_probabilities_for_persona(persona_type):  
    """  
    Επιστρέφει τις πιθανότητες για τις απαντήσεις [0, 1, 2, 3, 4]  
    βασισμένες σε Κανονική Κατανομή (Gaussian).  
    """  
    if persona_type == 'Minimal':  
        mu, sigma = 0.5, 0.8 # Κέντρο στο 0 (Υγιής)  
    elif persona_type == 'Mild':  
        mu, sigma = 1.5, 1.0  
    elif persona_type == 'Moderate':  
        mu, sigma = 2.5, 1.0  
    elif persona_type == 'Severe':  
        mu, sigma = 3.5, 0.8 # Κέντρο στο 4 (Σοβαρή)  
  
    # Υπολογισμός PDF και κανονικοποίηση ώστε το άθροισμα να είναι 1  
    probs = [norm.pdf(x, mu, sigma) for x in range(5)]  
    total = sum(probs)  
    return [p / total for p in probs]
```

Κώδικας 4.1: Υπολογισμός πιθανοτήτων Gaussian Personas (από `data_generator.py`)

4.2.6 Μηχανική Προτροπών (Prompt Engineering)

Για τη μετατροπή των αριθμητικών προφίλ σε φυσική γλώσσα, χρησιμοποιήθηκε το Google Gemini API. Εφαρμόστηκε η τεχνική Context Injection, παρέχοντας στο μοντέλο περιορισμούς (Constraints) για να μειωθούν οι παραισθήσεις (hallucinations), όπως: «Συμπτώματα ΑΠΟΝΤΑ: Αυτοκτονικός Ιδεασμός (Score 0)».

Η πρόσβαση στο Gemini εξασφαλίστηκε μέσω του Google AI Studio για αυθεντικοποίηση και λήψη API Keys.



Εικόνα 4.2: Διεπαφή του Google AI Studio για τη διαχείριση API Keys

4.3 Προεπεξεργασία και Εκπαίδευση

Πριν την εισαγωγή τους στο μοντέλο, τα δεδομένα κειμένου υποβάλλονται σε μια σειρά διαδικασιών εξυγίανσης και μετασχηματισμού. Στόχος είναι η μείωση του θορύβου και η βελτιστοποίηση της εισόδου για την αρχιτεκτονική του BERT.

4.3.α Καθαρισμός και Tokenization

Το πρώτο στάδιο αφορά την αφαίρεση χαρακτήρων που δεν προσθέτουν σημασιολογική αξία. Χρησιμοποιώντας κανονικές εκφράσεις (Regular Expressions), εφαρμόζουμε τους εξής κανόνες:

- Αφαίρεση θορύβου: Διαγράφονται emojis, ειδικά σύμβολα και χαρακτήρες ελέγχου, διατηρώντας μόνο τους αλφαριθμητικούς χαρακτήρες και τα βασικά σημεία στίξης.
- Κανονικοποίηση κενών: Πολλαπλά κενά ή αλλαγές γραμμής αντικαθίστανται από ένα ενιαίο κενό διάστημα.
- Προσομοίωση πραγματικών συνθηκών: Κατά τη φάση της εκπαίδευσης, εφαρμόζεται δυναμική αφαίρεση τόνων. Αυτό εκπαιδεύει το μοντέλο να είναι ανθεκτικό σε κείμενα που προέρχονται από κινητές συσκευές, όπου η χρήση τόνων συχνά παραλείπεται.

```
import re

def clean_text_for_bert(text):
    """
    Καθαρίζει το κείμενο για το BERT:
    1. Αφαιρεί πολλαπλά κενά (whitespace normalization).
    2. Αφαιρεί χαρακτήρες που δεν είναι λέξεις ή σημεία στίξης.
    """
    # Αντικατάσταση πολλαπλών κενών/newlines με ένα κενό
    text = re.sub(r'\s+', ' ', text).strip()

    # Διατήρηση μόνο αλφαριθμητικών και βασικής στίξης
    text = re.sub(r'^[\w\s.,!?!]', '', text)

    return text
```

Κώδικας 4.2: Συνάρτηση εξυγίανσης κειμένου με χρήση Regex (από `data_toolkit.py`).

4.3.2 Tokenization (Τεμαχισμός)

Μετά τον καθαρισμό, το κείμενο μετατρέπεται σε tokens χρησιμοποιώντας τον αλγόριθμο WordPiece, όπως αυτός έχει υλοποιηθεί στο Greek BERT. Η επιλογή του WordPiece είναι κρίσιμη για την ελληνική γλώσσα, καθώς διασπά τις λέξεις σε υπομονάδες (subwords). Αυτή η προσέγγιση επιτρέπει στο μοντέλο να αναγνωρίζει ρίζες και καταλήξεις, διαχειριζόμενο αποτελεσματικά σπάνιες ή άγνωστες λέξεις που δεν υπάρχουν στο αρχικό λεξιλόγιο εκπαίδευσης.

4.3.γ Στρατηγική Εκπαίδευσης (Multi-Task Learning)

Αντί για απλή ταξινόμηση, το μοντέλο εκπαιδεύτηκε να προβλέπει πολλαπλούς στόχους ταυτόχρονα (Multi-Task Regression):

- Για την κλίμακα Beck, το μοντέλο καλείται να προβλέψει 21 ξεχωριστά scores, τα οποία αντιστοιχούν στις ερωτήσεις του ερωτηματολογίου.
- Για την κλίμακα SCL-90, προβλέπονται αντίστοιχα 39 scores.

Η υλοποίηση πραγματοποιήθηκε μέσω της κατάλληλης παραμετροποίησης της κεφαλής εξόδου (Output Head) του BERT. Ορίστηκε η παράμετρος `num_labels` σε 21 και 39 αντίστοιχα, επιτρέποντας στο μοντέλο να παράγει πολλαπλές εξόδους. Ως συνάρτηση κόστους για την εκπαίδευση επιλέχθηκε το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error - MSE), το οποίο εφαρμόζεται ξεχωριστά σε κάθε έξοδο.

4.4 Αρχιτεκτονική Εξυπηρετητή (Backend Implementation)

Ο εξυπηρετητής αναπτύχθηκε με FastAPI. Η επιλογή αυτή έγινε λόγω της ασύγχρονης φύσης του (`async/await`), που επιτρέπει την ταυτόχρονη επεξεργασία αιτημάτων.

4.4.α Υβριδική Μηχανή Συμπερασμού

Στο αρχείο `Doctor.py` υλοποιείται η λογική που συνδυάζει τα αποτελέσματα. Το επίπεδο ελέγχου (Safety Layer με Regex) εκτελείται πρώτο για να εντοπίσει λέξεις-κλειδιά που ενδέχεται να υποδεικνύουν υψηλό κίνδυνο (π.χ. αυτοκτονία). Στη συνέχεια, εκτελείται το επίπεδο AI (BERT) για να ποσοτικοποιήσει τη συναισθηματική φόρτιση.

Στον Κώδικα 4.3 παρουσιάζεται η συνάρτηση `analyze_text`, η οποία φορτώνει τους τανυστές (tensors) και εκτελεί την πρόβλεψη.

```
def analyze_patient(scl_tok, scl_mod, beck_tok, beck_mod, anx_text, dep_text):
    # 1. Προετοιμασία εισόδου για SCL-90 (Άγχος)
    inputs_scl = scl_tok(anx_text, return_tensors="pt", truncation=True, max_length=512)

    # 2. Προετοιμασία εισόδου για Beck (Κατάθλιψη)
    inputs_beck = beck_tok(dep_text, return_tensors="pt", truncation=True, max_length=512)

    with torch.no_grad(): # Απενεργοποίηση gradients για ταχύτητα
        logits_scl = scl_mod(**inputs_scl).logits
        logits_beck = beck_mod(**inputs_beck).logits

    # 3. Μετατροπή Logits σε Scores και Clipping στο εύρος [0, 4]
    scl_items = np.clip(logits_scl[0].numpy(), 0.0, 4.0)
    beck_items = np.clip(logits_beck[0].numpy(), 0.0, 4.0)

    return scl_items, beck_items
```

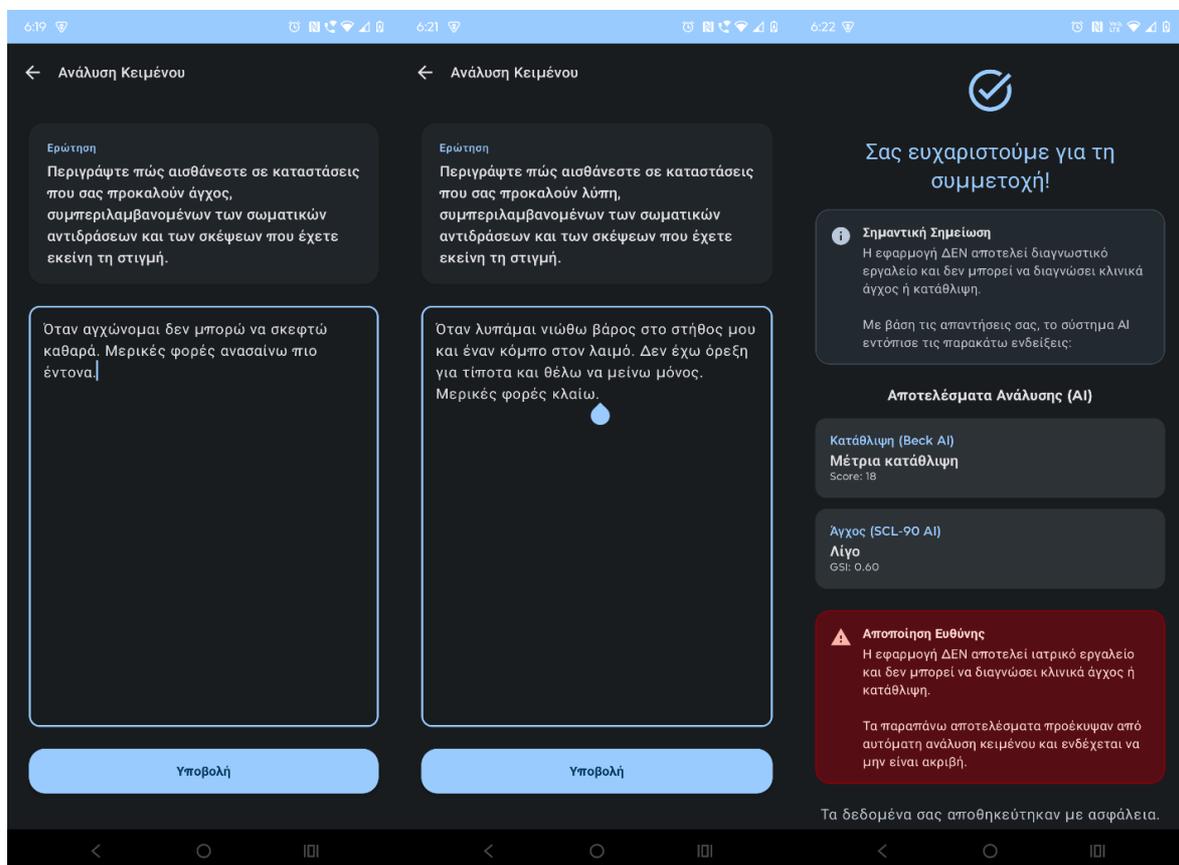
Κώδικας 4.3: Η συνάρτηση πρόβλεψης στο Backend (από το `Doctor.py`)

4.5 Ανάπτυξη Εφαρμογής Android (Frontend)

Η διεπαφή με τον χρήστη υλοποιήθηκε σε Kotlin με χρήση του Jetpack Compose. Η εφαρμογή σχεδιάστηκε με στόχο την προστασία της ιδιωτικότητας (Privacy by Design). Δεν απαιτεί εγγραφή χρήστη και τα δεδομένα αποθηκεύονται προσωρινά.

4.5.α Επικοινωνία με το Backend

Η εφαρμογή χρησιμοποιεί τη βιβλιοθήκη Retrofit για να στείλει τα κείμενα στον Server μέσω ασφαλούς σύνδεσης (HTTPS/Zrok Tunnel). Τα δεδομένα μορφοποιούνται σε JSON αντικείμενα (NetworkDTOs) πριν την αποστολή.



Εικόνες 4.3: Οθόνες της εφαρμογής Android. Αριστερά: Εισαγωγή κειμένου. Δεξιά: Παρουσίαση αποτελεσμάτων.

4.6 Διασφάλιση Ποιότητας και GDPR

Η αρχιτεκτονική επιδιώκει τη συμμόρφωση με τον ΓΚΠΔ (GDPR) μέσω της τεχνικής της Ψευδωνυμοποίησης. Κάθε χρήστης λαμβάνει ένα τυχαίο ID (UUID) και καμία προσωπική πληροφορία (Όνομα, Email) δεν αποστέλλεται στον Server. Επιπλέον, το σύστημα περιλαμβάνει μηχανισμούς Rate Limiting (μέσω της βιβλιοθήκης slowapi στον Server) για την προστασία από επιθέσεις DDoS, επιτρέποντας έως 20 αιτήματα ανά λεπτό.

4.7 Δήλωση Ηθικής (Ethics Statement)

Η παρούσα ενότητα αποσαφηνίζει τις ηθικές διαστάσεις της εργασίας.

Δεδομένα και Ιδιωτικότητα: Η εργασία δεν συλλέγει, δεν επεξεργάζεται και δεν αποθηκεύει δεδομένα πραγματικών ασθενών ή χρηστών. Το σύνολο των δεδομένων που χρησιμοποιήθηκαν για την εκπαίδευση του μοντέλου είναι συνθετικά, παραγόμενα αλγοριθμικά μέσω της μεθοδολογίας Gaussian Personas που περιγράφηκε στην Ενότητα 4.2. Κανένα δεδομένο δεν αντιστοιχεί σε πραγματικό πρόσωπο.

Έγκριση Ηθικής: Δεδομένου ότι η εργασία δεν περιλαμβάνει έρευνα σε ανθρώπινα υποκείμενα και δεν χρησιμοποιεί πραγματικά δεδομένα ασθενών, δεν απαιτήθηκε έγκριση από Επιτροπή Ηθικής και Δεοντολογίας (IRB approval).

Περιορισμοί Χρήσης: Το σύστημα αποτελεί τεχνική απόδειξη σκοπιμότητας (proof-of-concept) για ερευνητικούς σκοπούς και δεν προορίζεται για κλινική χρήση. Δεν υποκαθιστά την κλινική αξιολόγηση από εξειδικευμένο επαγγελματία ψυχικής υγείας. Οποιαδήποτε μελλοντική εφαρμογή σε πραγματικούς χρήστες θα απαιτούσε έγκριση από Επιτροπή Ηθικής, ενημερωμένη συγκατάθεση των συμμετεχόντων, και επικύρωση σε κλινικά δεδομένα.

Ασφάλεια: Το σύστημα περιλαμβάνει ευρετικούς μηχανισμούς για τον εντοπισμό ενδειξων κινδύνου, οι οποίοι όμως δεν έχουν αξιολογηθεί σε πραγματικές συνθήκες. Η αποτελεσματικότητά τους παραμένει υποθετική μέχρι την επικύρωση με κλινικά δεδομένα.

Κεφάλαιο 5: ΤΕΧΝΙΚΗ ΥΛΟΠΟΙΗΣΗ

5.1 Αρχιτεκτονική Εφαρμογής

Η αρχιτεκτονική ακολουθεί το πρότυπο Client-Server, διαχωρίζοντας τη διεπαφή χρήστη (Android) από την επεξεργαστική λογική (Backend). Αυτή η επιλογή διασφαλίζει απόδοση σε κινητές συσκευές περιορισμένων πόρων, καθώς η εκτέλεση μοντέλων Transformer απαιτεί σημαντική ισχύ GPU/CPU που δεν είναι διαθέσιμη στα περισσότερα smartphones.

5.1.α Δυναμική Επιλογή Συσκευής

Ο κώδικας είναι hardware-agnostic με δυναμική ανίχνευση πόρων. Κατά την εκκίνηση, το σύστημα ελέγχει για GPU μέσω `torch.cuda.is_available()`. Αν εντοπιστεί συμβατή κάρτα, το μοντέλο μεταφέρεται στη VRAM. Διαφορετικά, μεταπίπτει σε CPU.

Ο Κώδικας 5.1 παρουσιάζει τον μηχανισμό αυτόματης επιλογής συσκευής:

```
import torch

# Δυναμική ανίχνευση υλικού (Hardware Detection)
# Αν υπάρχει CUDA (NVIDIA GPU), χρησιμοποιείται για επιτάχυνση.
# Αλλιώς, γίνεται fallback σε CPU.
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
print(f"Using device: {device}")

# Μεταφορά του μοντέλου στη μνήμη της συσκευής (VRAM ή RAM)
model.to(device)
```

Κώδικας 5.1: Μηχανισμός δυναμικής επιλογής συσκευής εκτέλεσης (από `train_model_beck.py` & `Doctor.py`)

Η επικοινωνία μεταξύ client και server γίνεται μέσω REST API με HTTPS. Στη φάση ανάπτυξης χρησιμοποιήθηκε υπηρεσία tunneling (zrok) για προσωρινή έκθεση του τοπικού API. Για παραγωγική χρήση απαιτούνται επιπλέον μέτρα: αυθεντικοποίηση (API keys/mTLS), περιορισμός IP/ACL, και logging/monitoring [32].

Η εκπαίδευση του BERT πραγματοποιήθηκε στην πλατφόρμα Google Colab με NVIDIA T4 GPU, μειώνοντας τον χρόνο εκπαίδευσης σε περίπου μισή ώρα.

Change runtime type

Runtime type

Python 3

Hardware accelerator ?

CPU T4 GPU H100 GPU A100 GPU
 L4 GPU v5e-1 TPU v6e-1 TPU

i Want access to premium GPUs? [Purchase additional compute units](#)

Runtime version ?

Latest (recommended)

Cancel Save

Εικόνα 5.1: Το περιβάλλον εκπαίδευσης στο Google Colab με ενεργοποιημένη επιτάχυνση GPU

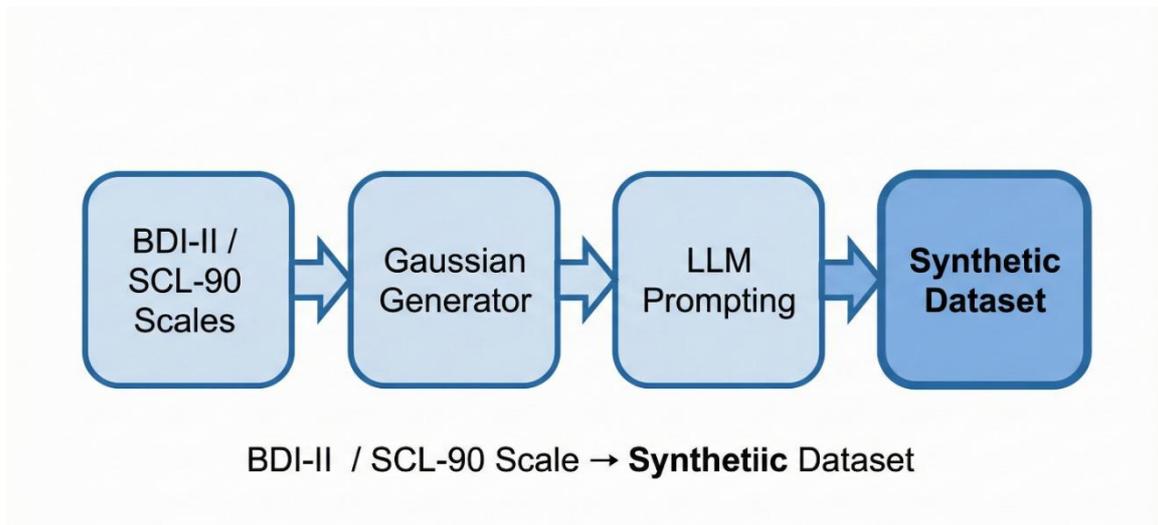
5.2 Υλοποίηση Backend (FastAPI & Python)

Το backend υλοποιήθηκε με FastAPI λόγω της υποστήριξης `async/await` και της ευκολίας δημιουργίας REST endpoints [34]. Η ασύγχρονη διαχείριση επιτρέπει εξυπηρέτηση πολλαπλών αιτημάτων χωρίς μπλοκάρισμα σε I/O-εντολές.

Σημειώνεται ότι το Global Interpreter Lock (GIL) της Python περιορίζει την παραλληλία σε single-process threads για CPU-bound εργασίες. Για βαριά PyTorch inference μπορούν να χρησιμοποιηθούν: `ThreadPoolExecutor/thread pools` για blocking κλήσεις [35], `multiprocessing/process pools` για πλήρη παράλληλη χρήση πολλαπλών πυρήνων [36], ή `worker processes (Celery/RQ)` για αποσύνδεση του inference από το web server [36][39].

5.2.α Μηχανισμός Παραγωγής Συνθετικών Δεδομένων (Synthetic Data Pipeline)

Για να αντιμετωπιστεί η έλλειψη ανοιχτών δεδομένων στα Ελληνικά (λόγω GDPR [23]), αναπτύχθηκε pipeline παραγωγής συνθετικών δεδομένων που βασίζεται σε Personas (Gaussian profiles). Η ροή της διαδικασίας παραγωγής των δεδομένων αυτών παρουσιάζεται στην Εικόνα 5.2.



Εικόνα 5.2: Ροή Παραγωγής Συνθετικών Δεδομένων

Σχήμα Ονοματοδοσίας ASCII-Safe

Για πρόληψη σφαλμάτων κωδικοποίησης κατά την αποθήκευση ελληνικών κειμένων σε CSV/βάσεις δεδομένων, κάθε ερώτηση αντιστοιχίζεται σε αλφαριθμητικό κλειδί στα Αγγλικά:

```

# config.py
# Αντιστοίχιση Ελληνικών ερωτήσεων σε ASCII keys
# για αποφυγή προβλημάτων encoding (utf-8) σε CSV/Databases.
SCL90_COLUMN_NAMES = {
    # I. ΣΩΜΑΤΟΠΟΙΗΣΗ (Somatization)
    'Q1': 'somatization_headaches',
    'Q4': 'somatization_faintness_dizziness',
    'Q12': 'somatization_heart_chest_pain',
    'Q27': 'somatization_low_back_pain',
    # ...
}
  
```

Κώδικας 5.2: Διαμόρφωση σχήματος δεδομένων ASCII-Safe (από config.py).

Αυτό εξασφαλίζει ότι τα αρχεία CSV ανοίγονται σε οποιοδήποτε λειτουργικό σύστημα χωρίς encoding issues και ο κώδικας παραμένει ευανάγνωστος.

Οι μέσες τιμές και τυπικές αποκλίσεις της βιβλιογραφίας μετασχηματίζονται σε persona-level προφίλ. Η πιθανότητα κάθε επιλογής (0-4) υπολογίζεται μέσω της CDF (Cumulative distribution function) της κανονικής κατανομής:

```
SCL90_PERSONA_CONFIGS = {
    'minimal': {'mean_mult': 0.3, 'sd_mult': 0.7},
    'mild':    {'mean_mult': 0.8, 'sd_mult': 1.0},
    'moderate': {'mean_mult': 1.5, 'sd_mult': 1.3},
    'severe':  {'mean_mult': 2.2, 'sd_mult': 1.6}
}

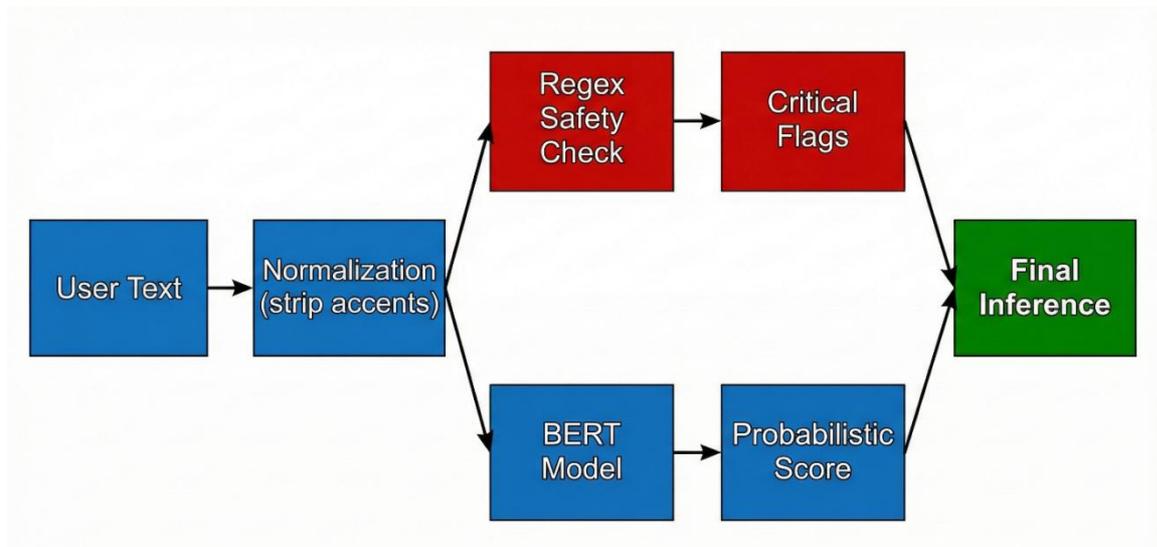
def get_probabilities_for_persona(base_mean, base_sd, scale_max, persona_c
onfig):
    pop_mean = base_mean * persona_config['mean_mult']
    pop_sd = base_sd * persona_config['sd_mult']
    boundaries = np.arange(-0.5, scale_max + 1.5, 1)
    z_scores = (boundaries - pop_mean) / pop_sd
    cdf_values = stats.norm.cdf(z_scores)
    probs = np.diff(cdf_values)
    return probs / probs.sum()
```

Κώδικας 5.3: Υπολογισμός Πιθανοτήτων Gaussian Persona (*data_generator.py*)

Η μέθοδος επιτυγχάνει ρεαλιστικότητα (τα συνθετικά δεδομένα ακολουθούν στατιστικές κατανομές παρόμοιες με τα κλινικά) και ποικιλομορφία (εκπαίδευση σε διαφορετικά επίπεδα σοβαρότητας).

5.2.β Υβριδική Μηχανή Συμπερασμού: Σύντηξη BERT και Ευρετικών Κανόνων (Hybrid Fusion)

Για την αξιολόγηση των κειμένων, το σύστημα υιοθετεί μια Υβριδική Αρχιτεκτονική (Hybrid Architecture). Η προσέγγιση αυτή συνδυάζει τη στοχαστική ικανότητα κατανόησης του μοντέλου BERT με έναν ευρετικό μηχανισμό ελέγχου βάσει κανόνων (Regex Safety Check), ο οποίος στοχεύει να λειτουργήσει ως επιπλέον επίπεδο ελέγχου για σημαντικές περιπτώσεις [33]. Η συνολική αρχιτεκτονική της υβριδικής μηχανής απεικονίζεται στην Εικόνα 5.3.



Εικόνα 5.3: Διάγραμμα Υβριδικής Μηχανής

Η υλοποίηση του backend απαιτεί προσεκτικό σχεδιασμό τόσο σε επίπεδο κώδικα όσο και σε επίπεδο υποδομής. Για τη διασφάλιση της ανωνυμίας και της συμμόρφωσης με το GDPR, κάθε συμμετέχων αποκτά ένα μοναδικό κρυπτογραφημένο αναγνωριστικό (hashed ID). Το αρχικό Android ID της συσκευής κρυπτογραφείται με SHA-256 και αποθηκεύεται μόνο το hash.

Ο μηχανισμός αυτός, μαζί με τη διαχείριση διπλότυπων εγγραφών, υλοποιείται στον Κώδικα 5.4:

```

# api/services/id_service.py
import hashlib
from typing import Set

def _load_used_identifiers() -> Set[str]:
    """Φόρτωση κρυπτογραφημένων IDs από αρχείο"""
    if not IDENTIFIERS_FILE.exists():
        return set()
    with open(IDENTIFIERS_FILE, 'r', encoding='utf-8') as f:
        return {line.strip() for line in f if line.strip()}

def register_identifier(identifier: str) -> str:
    """
    Καταγραφή νέου ID με SHA-256 hashing
    Παράδειγμα: "abc123" -> "9712168c..." (64 chars)
    """
    if not identifier:
        raise ValueError("Identifier cannot be empty")

    with identifier_lock:
        used = _load_used_identifiers()

        if identifier in used:
            return "already_exists"

    # Αποθήκευση με append για atomicity
    with open(IDENTIFIERS_FILE, 'a', encoding='utf-8') as f:
        f.write(identifier + '\n')
  
```

```
logger.info(f"Registered new ID: {identifier[:8]}...")
return "success".
```

Κώδικας 5.4: Μηχανισμός κρυπτογράφησης και ελέγχου αναγνωριστικών για διασφάλιση GDPR compliance (από id_service.py).

Κάθε συμμετέχων και συμμετέχουσα αποκτά ένα μοναδικό κρυπτογραφημένο αναγνωριστικό που αποθηκεύεται στο αρχείο used_identifiers.txt, αποτρέποντας διπλές συμμετοχές χωρίς να αποθηκεύονται προσωπικά δεδομένα. Έτσι επιτυγχάνεται ανωνυμία (το πρωτότυπο ID δεν αποθηκεύεται) και μοναδικότητα (κάθε συσκευή συμμετέχει μόνο μία φορά).

Διαχείριση Κύκλου Ζωής Server

Κατά την εκκίνηση, το σύστημα φορτώνει τα μοντέλα BERT στη μνήμη και εκτελεί προθέρμανση (warm-up) για να μειώσει το latency των πρώτων αιτημάτων.

```
# api/server.py
from contextlib import asynccontextmanager

@asynccontextmanager
async def lifespan(app: FastAPI):
    """
    Lifespan manager: Εκτελείται κατά την εκκίνηση και τερματισμό του server.
    Φορτώνει τα BERT μοντέλα στη μνήμη και εκτελεί warm-up inference.
    """
    # --- STARTUP LOGIC ---
    print("⌚ Initializing Hybrid AI System...")
    try:
        result = load_system() # Φόρτωση BERT models από Doctor.py
        if result:
            (model_state.s_tok, model_state.s_mod, model_state.s_scal,
             model_state.b_tok, model_state.b_mod, model_state.b_scal) = result

        if model_state.s_tok:
            print("Server READY!")

            # Aggressive warm-up για βελτιστοποίηση απόδοσης
            print("Starting aggressive warm-up sequence...")
            warm_up_texts = [
                ("Νιώθω νευρικότητα και φόβο", "Έχω λύπη και απώλεια ενδιαφέροντος")
            ]

            for idx, (anx, dep) in enumerate(warm_up_texts, 1):
                print(f" ...Warm-up pass {idx}/{len(warm_up_texts)}")
                analyze_patient(
                    model_state.s_tok, model_state.s_mod, model_state.s_sc
```

```

al,
        model_state.b_tok, model_state.b_mod, model_state.b_sc
al,
        anx, dep
    )
    print("☑ Models warmed up and stabilized!")
else:
    print("✗ FAILED to load models.")
except Exception as e:
    print(f"✗ Startup Error: {e}")

yield # Server runs here

print("Server shutting down...")

```

Κώδικας 5.4: Διαχείριση κύκλου ζωής του FastAPI server (από `server.py`).

Endpoint Ανάλυσης

Το κεντρικό endpoint του backend είναι το `/analyze`, το οποίο δέχεται τα κείμενα του χρήστη, εκτελεί την πρόβλεψη και επιστρέφει τα αποτελέσματα.

```

# api/routers/analysis.py
@router.post("/analyze")
@limiter.limit("20/minute")
async def analyze_symptoms(request: Request, data: AnalysisInput):
    """
    Endpoint ανάλυσης κειμένου με BERT.
    Δέχεται anxiety_text και depression_text, επιστρέφει predictions.
    """
    if not model_state.s_tok:
        raise HTTPException(status_code=503, detail="Models not initialized")

    try:
        # Εκτέλεση inference με τα φορτωμένα μοντέλα
        scl_preds, beck_preds, scl_items, beck_items = await loop.run_in_executor(
            executor, analyze_patient,
            model_state.s_tok, model_state.s_mod, model_state.s_scal,
            model_state.b_tok, model_state.b_mod, model_state.b_scal,
            data.anxiety_text, data.depression_text
        )

        # Logging για debugging και monitoring
        logger.info(f"Analysis success: SCL GSI={scl_preds[4]:.2f}, Beck={beck_preds[0]:.1f}")

        return {
            "scl90": {
                "gsi": float(scl_preds[4]),
                "pst": float(scl_preds[5]),
            }
        }
    
```

```

        "items": [float(x) for x in scl_items] # 39 predictions
    },
    "beck": {
        "total_score": float(beck_preds[0]),
        "items": [float(x) for x in beck_items] # 21 predictions
    }
}
except Exception as e:
    logger.error(f"Inference Error: {e}", exc_info=True)
    raise HTTPException(status_code=500, detail="Analysis failed")

```

Κώδικας 5.6: Endpoint ανάλυσης κειμένου με real-time inference (από analysis.py).

The screenshot shows a Windows File Explorer window titled 'emotion_app_results'. The address bar shows the path 'Thesis_Complete_Project > emotion_app_results'. The file list includes:

Name	Date modified	Type	Size
.gitkeep	1/27/2026 7:17 PM	GITKEEP File	0 KB
all_participants.xlsx	1/28/2026 11:51 AM	Φύλλο εργασίας τ...	8 KB
participant_20260127_231115.xlsx	1/27/2026 11:11 PM	Φύλλο εργασίας τ...	13 KB
participant_20260127_234643.xlsx	1/27/2026 11:46 PM	Φύλλο εργασίας τ...	13 KB
participant_20260128_000415.xlsx	1/28/2026 12:04 AM	Φύλλο εργασίας τ...	13 KB
participant_20260128_115140.xlsx	1/28/2026 11:51 AM	Φύλλο εργασίας τ...	13 KB
used_identifiers.txt	1/28/2026 11:51 AM	Text Document	1 KB

An inset window titled 'used_identifiers.txt' shows the following content:

```

9712168c36341fb9
f95b1ed66b1a828b
8f290e8bbd9bc826

```

The status bar at the bottom of the text editor shows: Ln 1, Col 1 | 51 characte | Plain te> | 100% | Windows | UTF-8

```

Thesis API Server - python ap x + v
THESES SERVER (REFACTORED) - V2.0.0
Modular Architecture with Vectorized Processing

INFO: Started server process [20608]
INFO: Waiting for application startup.
  Initializing Hybrid AI System...
INFO:Thesis_Doctor:Loading Hybrid System (Item-Level)...
INFO:Thesis_Doctor:Models loaded successfully!
  Server READY!
  Starting aggressive warm-up sequence...
  ..Warm-up pass 1/2
  ..Warm-up pass 2/2
  Models warmed up and stabilized!
INFO: Application startup complete.
INFO: Uvicorn running on http://0.0.0.0:8000 (Press CTRL+C to quit)
INFO: 10.32.1.241:0 - "GET /check-id?id=8f290e8bbd9bc826 HTTP/1.1" 200 OK
INFO: 10.32.0.136:0 - "GET /check-id?id=8f290e8bbd9bc826 HTTP/1.1" 200 OK
INFO:Thesis_Server:Analysis success: SCL GSI=1.64, Beck=0.0
INFO: 10.32.0.136:0 - "POST /analyze HTTP/1.1" 200 OK
INFO:Thesis_Server:Analysis success: SCL GSI=0.00, Beck=14.7
INFO: 10.32.1.241:0 - "POST /analyze HTTP/1.1" 200 OK
INFO:Thesis_Server:Detailed Excel saved: C:\Users\sba\k\Desktop\Thesis_Complete_Project\emotion_app_results\participant_20260128_115140.xlsx
INFO:Thesis_Server:Aggregated Excel updated with INTERLEAVED Manual & AI data (vectorized).
INFO:Thesis_Server:Results saved for timestamp: 20260128_115140
INFO: 10.32.1.241:0 - "POST /submit-results HTTP/1.1" 200 OK
INFO:Thesis_Server:Registered new ID: 8f290e8bbd9bc826
INFO: 10.32.1.241:0 - "POST /register-id HTTP/1.1" 200 OK
INFO: 10.32.2.131:0 - "GET /check-id?id=8f290e8bbd9bc826 HTTP/1.1" 200 OK

Zrok Tunnel - python "C:\Use x + v
PUBLIC URL FOUND!
https://pa9riish06zn.share.zrok.io
=====
[4/4] Updating Android configuration...
Updated NetworkService.kt with new URL
-----
TUNNEL ACTIVE - DO NOT CLOSE THIS WINDOW
-----
Press Ctrl+C to stop the tunnel

[zrok]: {"file": "/home/runner/work/zrok/zrok/cmd/zrok/sharePublic.go:298", "func": "main.(*sharePublicCommand).shareLocal", "level": "info", "msg": "[ ] -\u003e GET /check-id?id=8f290e8bbd9bc826", "time": "2026-01-28T11:45:05.290Z"}
[zrok]: {"file": "/home/runner/work/zrok/zrok/cmd/zrok/sharePublic.go:298", "func": "main.(*sharePublicCommand).shareLocal", "level": "info", "msg": "[ ] -\u003e GET /check-id?id=8f290e8bbd9bc826", "time": "2026-01-28T11:46:06.432Z"}
[zrok]: {"file": "/home/runner/work/zrok/zrok/cmd/zrok/sharePublic.go:298", "func": "main.(*sharePublicCommand).shareLocal", "level": "info", "msg": "[ ] -\u003e POST /analyze", "time": "2026-01-28T11:47:15.958Z"}
[zrok]: {"file": "/home/runner/work/zrok/zrok/cmd/zrok/sharePublic.go:298", "func": "main.(*sharePublicCommand).shareLocal", "level": "info", "msg": "[ ] -\u003e POST /analyze", "time": "2026-01-28T11:48:41.041Z"}
[zrok]: {"file": "/home/runner/work/zrok/zrok/cmd/zrok/sharePublic.go:298", "func": "main.(*sharePublicCommand).shareLocal", "level": "info", "msg": "[ ] -\u003e POST /submit-results", "time": "2026-01-28T11:51:40.045Z"}
[zrok]: {"file": "/home/runner/work/zrok/zrok/cmd/zrok/sharePublic.go:298", "func": "main.(*sharePublicCommand).shareLocal", "level": "info", "msg": "[ ] -\u003e POST /register-id", "time": "2026-01-28T11:51:41.264Z"}
[zrok]: {"file": "/home/runner/work/zrok/zrok/cmd/zrok/sharePublic.go:298", "func": "main.(*sharePublicCommand).shareLocal", "level": "info", "msg": "[ ] -\u003e GET /check-id?id=8f290e8bbd9bc826", "time": "2026-01-28T11:52:04.714Z"}

```

Εικόνα 5.4: Υποδομή και λειτουργία Backend.

Πρώτη Εικόνα: Φάκελος αποτελεσμάτων με αυτόματη δημιουργία αρχείων Excel ανά συμμετέχοντα.

Δεύτερη Εικόνα: Κρυπτογραφημένα αναγνωριστικά (UUID hashes) για διασφάλιση ανωνυμίας (GDPR compliance).

Τρίτη Εικόνα: Console logs εκκίνησης server με φόρτωση μοντέλων BERT και επιτυχή επεξεργασία αιτημάτων (HTTP 200).

Τέταρτη Εικόνα: Zrok tunnel για remote testing κατά την ανάπτυξη.

Zrok Tunnel για Development

Για testing σε πραγματικές συσκευές Android κατά την ανάπτυξη, χρησιμοποιήθηκε το Zrok για δημιουργία ασφαλούς tunnel:

```

# Development only - Zrok tunnel για remote testing
# Εκτελείται σε ξεχωριστό terminal πριν την εκκίνηση του server

# 1. Εγκατάσταση zrok (one-time)
# pip install zrok

# 2. Εκκίνηση tunnel που προωθεί το τοπικό port 8000
# zrok share public localhost:8000

```

```

# Η εντολή επιστρέφει public URL:
# https://pa9riish06zn.share.zrok.io

# 3. Αυτόματη ενημέρωση Android configuration
# Το zrok script ενημερώνει το NetworkService.kt:
#
# // NetworkService.kt
# private val serverUrl = "https://pa9riish06zn.share.zrok.io"

# 4. Τα αιτήματα από το Android περνούν μέσω του tunnel:
# Android -> Zrok Tunnel -> Localhost:8000 -> FastAPI

# ΣΗΜΕΙΩΣΗ: Για παραγωγή απαιτείται:
# - Αυθεντικοποίηση (API keys / JWT)
# - Rate limiting (συμπεριλαμβάνεται: 20 req/min)
# - Monitoring και Logging (υλοποιημένο)

```

Κώδικας 5.7: Ρύθμιση Zrok tunnel για remote testing κατά την ανάπτυξη.

Το tunnel επιτρέπει στην εφαρμογή Android να επικοινωνεί με τον τοπικό server μέσω ενός δημόσιου HTTPS URL, χωρίς να απαιτείται cloud deployment. Τα logs της Εικόνας 5.4 (Τρίτη Εικόνα) δείχνουν τα αιτήματα που περνούν μέσω του tunnel (GET /check-id, POST /analyze, POST /submit-results).

Επίπεδο Ελέγχου με Regular Expressions

Η διαδικασία αξιολόγησης (Inference Pipeline) περιλαμβάνει, πριν το νευρωνικό δίκτυο, ένα επίπεδο ελέγχου με Regular Expressions. Αυτό λειτουργεί ως δικλείδα ασφαλείας για λέξεις-κλειδιά υψηλού κινδύνου, λαμβάνοντας υπόψη και την άρνηση (negation).

```

def detect_symptoms_advanced(text, mapping_dict):
    detected_items = set()
    text_normalized = strip_accents(text.lower())
    for keyword_pattern, item_description in mapping_dict.items():
        pattern_clean = strip_accents(keyword_pattern.lower())
        pattern = re.compile(pattern_clean, re.IGNORECASE | re.UNICODE)
        matches = list(pattern.finditer(text_normalized))
        for match in matches:
            context = text_normalized[max(0, match.start() - 35):match.start()]

            has_negation = any(neg in context for neg in NEGATIONS)
            if not has_negation:
                detected_items.add(item_description)
    return sorted(list(detected_items))

```

Κώδικας 5.8: Αλγόριθμος εντοπισμού συμπτωμάτων (Doctor.py)

Η λειτουργία περιλαμβάνει:

- Κανονικοποίηση: Αφαίρεση τόνων και μετατροπή σε πεζά ώστε οι λέξεις-κλειδιά να εντοπίζονται ανεξαρτήτως τρόπου γραφής
- Ευρετικός έλεγχος: Ανίχνευση ριζών λέξεων που σχετίζονται με κίνδυνο (π.χ. «αυτοκτον»)
- Έλεγχος άρνησης: Εξέταση παραθύρου χαρακτήρων πριν τη λέξη-στόχο ώστε φράσεις όπως «δεν θέλω να πεθάνω» να μην ενεργοποιούν λανθασμένα την ένδειξη κινδύνου

BERT Regression Layer

Μετά την ολοκλήρωση του καθαρισμού με Regex, το κείμενο διοχετεύεται στο Greek BERT [20][18]. Σε αντίθεση με τους κανόνες Regex που λειτουργούν δυαδικά (εντοπισμός ή μη), το BERT ποσοτικοποιεί τη συναισθηματική βαρύτητα, αποδίδοντας μια συνεχή τιμή στο διάστημα $[0, 4]$ για κάθε στοιχείο των κλιμάκων.

Ωστόσο, επειδή το μοντέλο λειτουργεί ως παλινδρομητής (regressor), υπάρχει ο κίνδυνος παραγωγής τιμών εκτός των επιτρεπτών ορίων (out-of-bound predictions). Για την αποφυγή αυτού του φαινομένου και τη διασφάλιση της εγκυρότητας των αποτελεσμάτων, εφαρμόζεται η τεχνική της περικοπής τιμών (value clipping).

Συγκεκριμένα, κάθε πρόβλεψη περιορίζεται αυστηρά στο εύρος $[0,4]$:

$$y_{\text{final}} = \max(0, \min(y_{\text{pred}}, 4))$$

Η υλοποίηση της διαδικασίας στο υποσύστημα παρατίθεται στον Κώδικα 5.9:

```
# Υπολογισμός Logits από το μοντέλο BERT
with torch.no_grad():
    logits_scl = scl_mod(**inputs_scl).logits
    logits_beck = beck_mod(**inputs_beck).logits

# Post-processing: Εφαρμογή Clipping στο εύρος [0.0, 4.0]
scl_items = np.clip(logits_scl[0].numpy(), 0.0, 4.0)
beck_items = np.clip(logits_beck[0].numpy(), 0.0, 4.0)
```

Κώδικας 5.9: Εφαρμογή ορίων (Clipping) στα αποτελέσματα του BERT (από Doctor.py).

Σύντηξη Αποτελεσμάτων

Η τελική εκτίμηση προκύπτει από συνδυασμό των δύο επιπέδων. Το Regex λειτουργεί ως ευρετικός μηχανισμός ελέγχου: εάν εντοπιστεί ένδειξη αυτοκτονικού ιδεασμού, λαμβάνεται υπόψη ανεξάρτητα από τη βαθμολογία BERT, στοχεύοντας σε υψηλή ευαισθησία σε θέματα ασφάλειας [33].

Σημείωση: Ο ισχυρισμός περί υψηλής ευαισθησίας δεν έχει επικυρωθεί σε πραγματικά κλινικά δεδομένα. Η πραγματική απόδοση σε περιπτώσεις κινδύνου παραμένει άγνωστη.

5.2.γ Εκπαίδευση Μοντέλου και Τεχνικές Επαύξησης Δεδομένων

Για την εκπαίδευση του Greek BERT εφαρμόστηκαν τεχνικές augmentation, συμπεριλαμβανομένων της αφαίρεσης τόνων, της προσομοίωσης greeklish, ορθογραφικών σφαλμάτων και fillers.

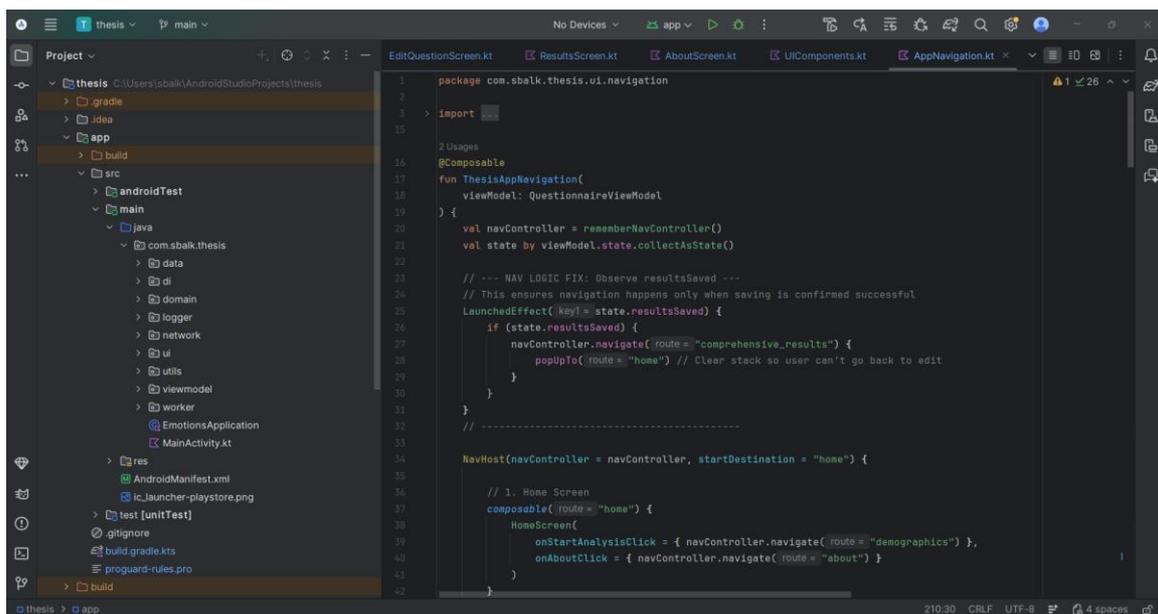
```
def augment_greek_text(text, intensity=0.3):
    if random.random() < intensity:
        text = text.replace(',', ' ')
    if random.random() < intensity * 0.4:
        text = remove_random_accent(text)
    if random.random() < intensity * 0.7:
        filler = random.choice(["Λοιπόν,", "Εεε,", "Νομίζω"])
        text = f"{filler} {text}"
    return text
```

Κώδικας 5.10: Τεχνική Data Augmentation για Ελληνικά (*data_toolkit.py*)

Για την ένταξη σε περιβάλλον παραγωγής, προτείνεται η υιοθέτηση τεχνικών συμπίεσης μοντέλου (όπως κβαντισμός, κλάδεμα και κβαντισμός μετά την εκπαίδευση), καθώς και η χρήση εξειδικευμένων περιβαλλόντων εκτέλεσης για κινητά (όπως τα PyTorch Mobile, ONNX Runtime Mobile και TFLite). Στόχος είναι η εκτέλεση της διαδικασίας συμπερασμού τοπικά στη συσκευή (on-device inference). Οι μέθοδοι αυτές επιτυγχάνουν μείωση της κατανάλωσης μνήμης και του χρόνου απόκρισης (latency), αν και συχνά συνεπάγονται μικρούς συμβιβασμούς ως προς την ακρίβεια του μοντέλου [37][29].

5.3 Εφαρμογή Android (Client)

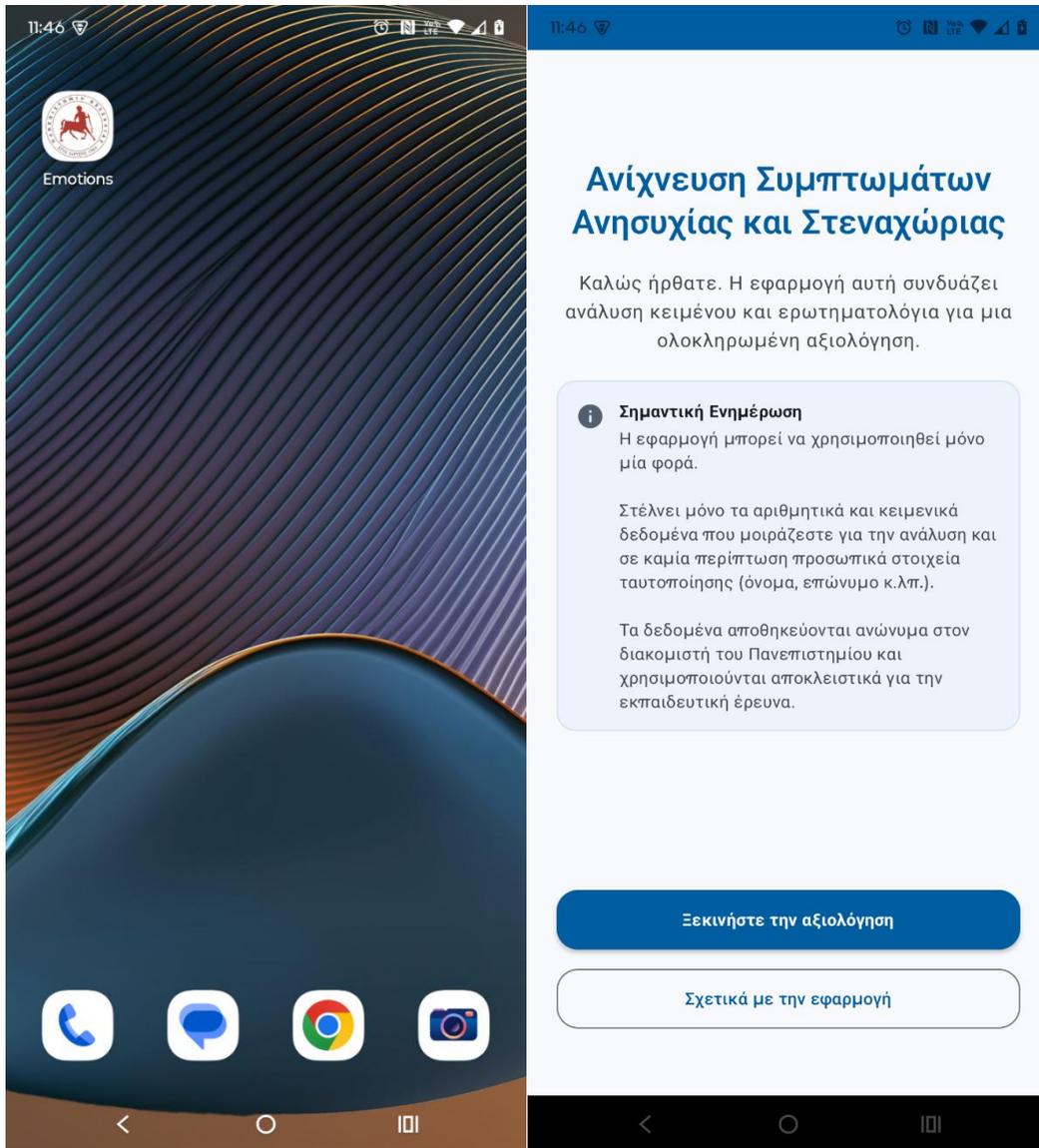
Η εφαρμογή αναπτύχθηκε στο Android Studio με Kotlin, την προτεινόμενη γλώσσα από την Google για Android development. Η διαχείριση εξαρτήσεων και η μεταγλώττιση έγιναν μέσω Gradle.



Εικόνα 5.5: Το περιβάλλον ανάπτυξης Android Studio και η δομή του project σε Kotlin

Η εφαρμογή χρησιμοποιεί Jetpack Compose και αρχιτεκτονική MVVM (Model–view–viewmodel), διαχωρίζοντας το UI από τη λογική για ευκολότερο testing και modularity.

Η αρχική οθόνη παρουσιάζει τον σκοπό του συστήματος και παρέχει οδηγίες χρήσης.



Εικόνα 5.6: Εκκίνηση της εφαρμογής. Αριστερά: Το εικονίδιο της Εφαρμογής Emotions στο μενού της συσκευής. Δεξιά: Η οθόνη υποδοχής με τον τίτλο της εφαρμογής και τις βασικές οδηγίες χρήσης.

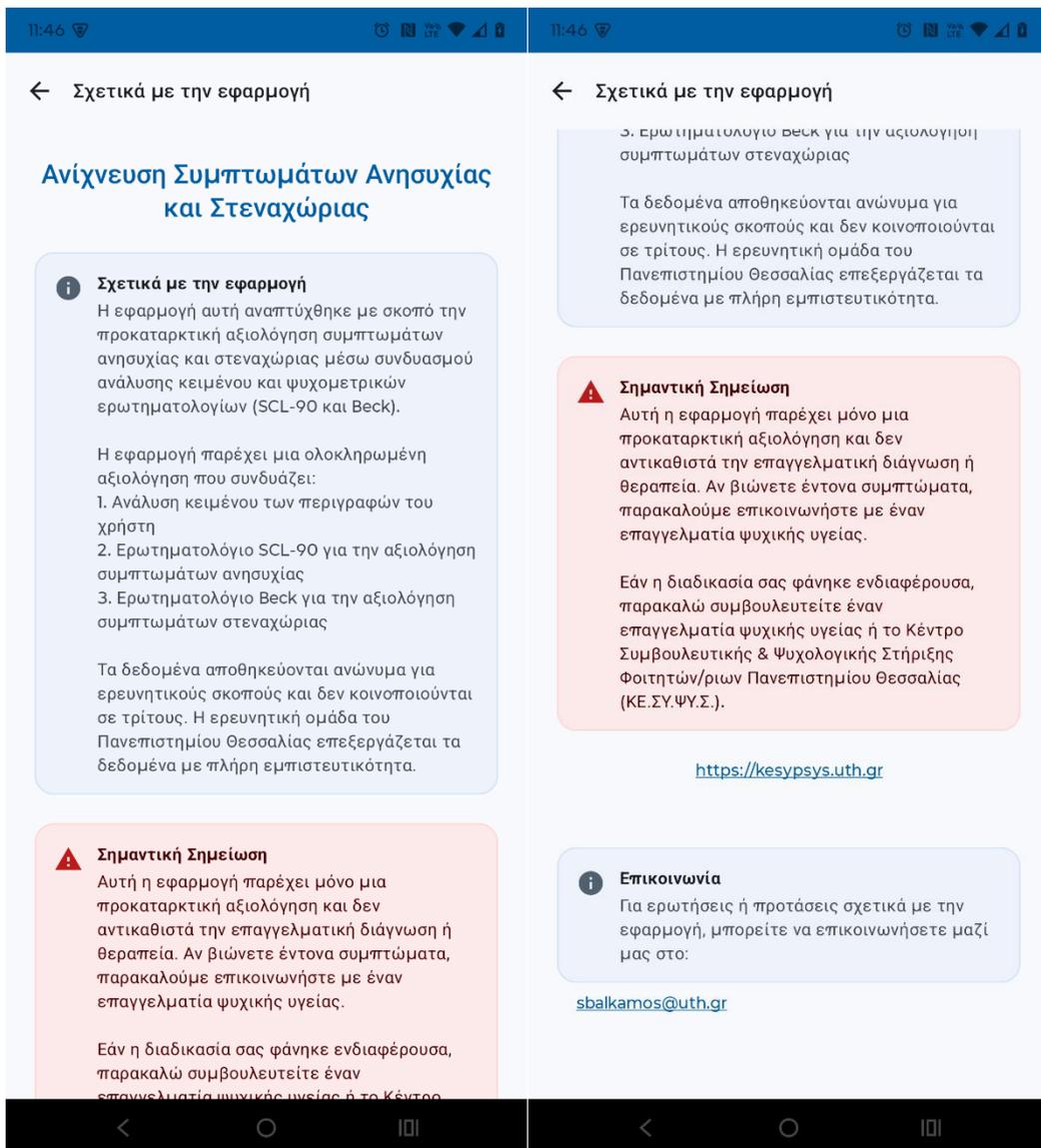
5.3.α Βασικά Χαρακτηριστικά:

- Τοπικός υπολογισμός: Άμεσοι υπολογισμοί SCL-90/Beck στη συσκευή (αριθμητικά, χωρίς heavy ML)
- Ασύγχρονη επικοινωνία: Kotlin Coroutines και StateFlow για non-blocking networking και responsive UI
- Privacy by Design: Χωρίς signup, με hashed/salted ANDROID_ID για ephemeral session identification, clear-on-exit memory hygiene και μη μόνιμη τοπική αποθήκευση [6][38]

Σημειώνεται ότι το ANDROID_ID έχει περιορισμούς και ιδιαιτερότητες (device resets, OEM behaviour). Συνιστάται salted hashing και σαφής πολιτική απορρήτου [38].

5.3.β Ενημερωμένη Συγκατάθεση

Η εφαρμογή περιλαμβάνει οθόνη πληροφοριών με ρητή δήλωση αποποίησης ευθύνης που τονίζει ότι δεν αποτελεί διαγνωστικό εργαλείο, στοιχεία επικοινωνίας με το ΚΕ.ΣΥ.ΨΥ.Σ. του Πανεπιστημίου Θεσσαλίας, και email του υπευθύνου ερευνητή.



Εικόνα 5.6: Οθόνες πληροφόρησης και ηθικής δέσμευσης. Αριστερά: Παρουσίαση του σκοπού και της μεθοδολογίας της εφαρμογής. Δεξιά: Ρητή δήλωση αποποίησης ευθύνης (Disclaimer) με κόκκινο πλαίσιο έμφασης, που επισημαίνει ότι η εφαρμογή δεν υποκαθιστά την κλινική διάγνωση.

5.3.γ Πολιτική Ασφάλειας Δικτύου

Η πολιτική ασφάλειας δικτύου της εφαρμογής ορίστηκε ρητά μέσω του αρχείου network_security_config.xml. Το αρχείο αυτό καθορίζει τους κανόνες εμπιστοσύνης για τα ψηφιακά πιστοποιητικά (SSL/TLS). Παρόλο που επιτρέπεται η κίνηση καθαρού κειμένου

(cleartext) για σκοπούς τοπικής αποσφαλμάτωσης (debugging), η παραγωγική επικοινωνία βασίζεται στα πιστοποιητικά συστήματος, διασφαλίζοντας την προστασία των δεδομένων κατά τη μεταφορά τους (data in transit).

```
<!-- network_security_config.xml -->
<?xml version="1.0" encoding="utf-8"?>
<network-security-config>
  <!-- Επιτρέπει cleartext μόνο για development/debugging -->
  <base-config cleartextTrafficPermitted="true">
    <trust-anchors>
      <!-- Εμπιστοσύνη μόνο στα πιστοποιητικά του συστήματος -->
      <certificates src="system" />
    </trust-anchors>
  </base-config>
</network-security-config>
```

Κώδικας 5.11: Ρυθμίσεις ασφάλειας δικτύου Android (από `network_security_config.xml`).

5.3.6 Συλλογή Δημογραφικών

Πριν την έναρξη της κύριας αξιολόγησης, η εφαρμογή συλλέγει βασικά δημογραφικά στοιχεία (φύλο και ηλικία) τα οποία χρησιμοποιούνται για την κατάλληλη προσαρμογή των ερωτηματολογίων. Συγκεκριμένα, το ερωτηματολόγιο Beck Depression Inventory II (BDI-II) έχει διαφορετικές εκδοχές για άνδρες και γυναίκες. Η διεπαφή είναι σχεδιασμένη για να είναι όσο το δυνατόν πιο απλή και κατανοητή, με μεγάλα κουμπιά επιλογής και αυτόματη επικύρωση των δεδομένων (validation) πριν την συνέχεια. Η εικόνα 5.7 απεικονίζει την οθόνη συλλογής δημογραφικών στοιχείων.

Εικόνα 5.7: Οθόνη εισαγωγής δημογραφικών στοιχείων. Ο χρήστης επιλέγει φύλο μέσω κουμπιών toggle και εισάγει την ηλικία με χρήση αριθμητικού πληκτρολογίου. Το σύστημα επικυρώνει ότι η ηλικία βρίσκεται στο εύρος 18-100 ετών πριν επιτρέψει τη συνέχεια.

5.3.ε Data Transfer Objects και Mapper Pattern

Για τη διασφάλιση της αρχιτεκτονικής καθαρότητας και την αποσύζευξη (decoupling) της διεπαφής χρήστη από το επίπεδο δικτύου, υιοθετήθηκε το πρότυπο των Data Transfer Objects (DTOs) σε συνδυασμό με έναν εξειδικευμένο Mapper. Αντί η εφαρμογή να αποστέλλει απευθείας τα εσωτερικά αντικείμενα κατάστασης (State Objects), χρησιμοποιείται η κλάση PayloadMapper η οποία μετασχηματίζει τα δεδομένα σε μια αυστηρά ορισμένη δομή JSON (NetworkSubmitRequest). Αυτή η πρακτική προστατεύει το API συμβόλαιο από αλλαγές στο UI και επιτρέπει την ασφαλή και τυποποιημένη ανταλλαγή δεδομένων.

```
// PayloadMapper.kt
fun mapToNetworkRequest(
    scl90: SCL90Results,
```

```

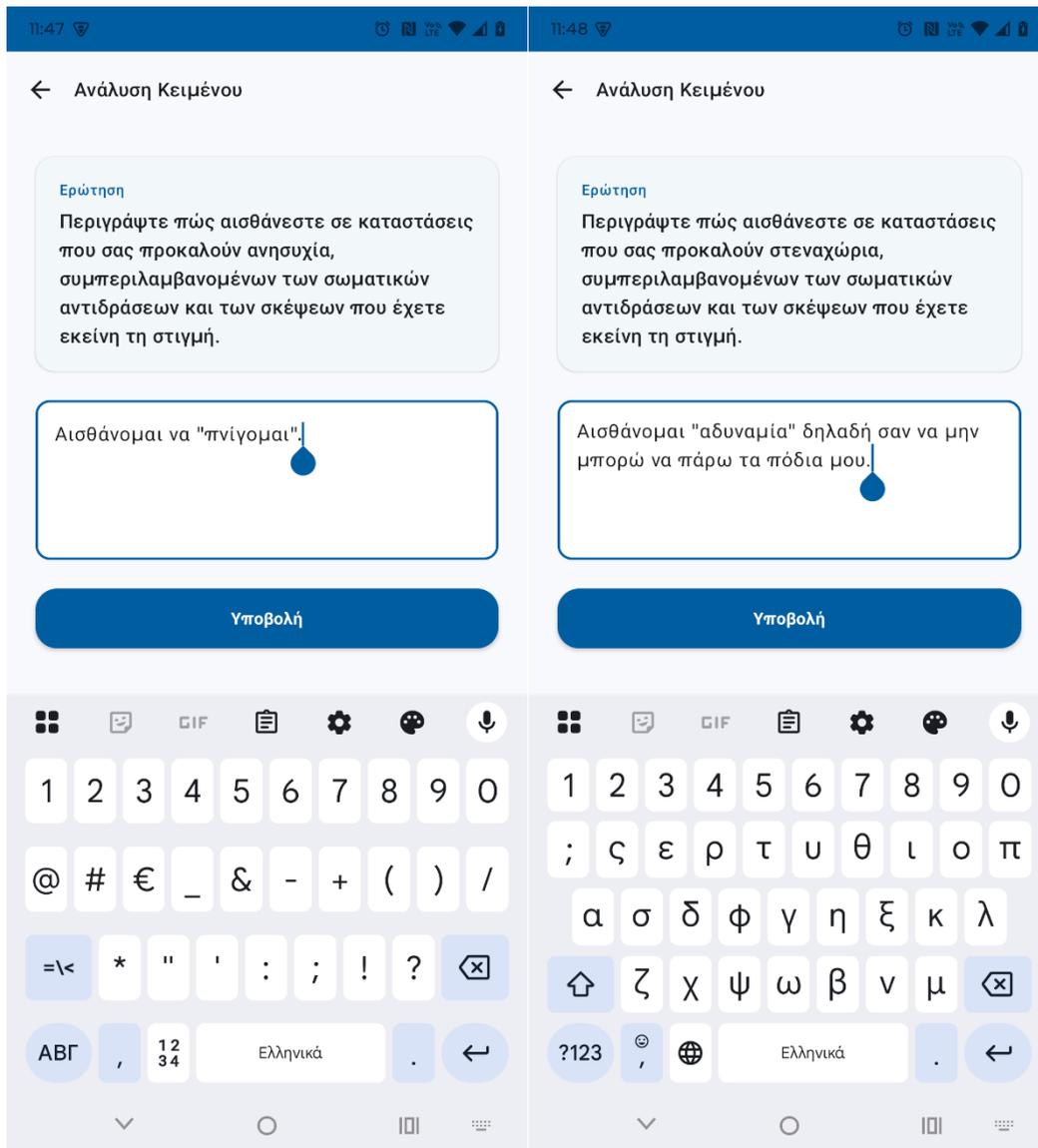
    beck: BeckResults
): NetworkSubmitRequest {
    val itemScores = mutableMapOf<String, Int>()
    scl90.detailedAnswers.forEach { (questionText, score) ->
        val id = sclIdMap[questionText]
        if (id != null) itemScores[id] = score
    }
    return NetworkSubmitRequest(
        scl90Results = NetworkScl90Data(..., itemScores),
        beckResults = NetworkBeckData(..., beckIndices)
    )
}

```

Κώδικας 5.12: Υλοποίηση του Mapper Pattern για τη διαχείριση αιτημάτων δικτύου (από PayloadMapper.kt).

5.3.στ Συλλογή Ελεύθερου Κειμένου

Μετά τη συλλογή των δημογραφικών στοιχείων, το σύστημα προχωρά στη συλλογή ελεύθερου κειμένου (free-text input). Ο/Η χρήστης καλείται να περιγράψει με δικά του/της λόγια τα συμπτώματα που βιώνει, σε δύο ξεχωριστές ερωτήσεις: μία για το άγχος και μία για τη στεναχώρια. Αυτή η προσέγγιση επιτρέπει στο μοντέλο BERT να αναλύσει το συναίσθημα και την έκφραση του χρήστη πέρα από τις δομημένες απαντήσεις των ψυχομετρικών εργαλείων. Το πεδίο κειμένου υποστηρίζει πολυγραμμική εισαγωγή (multiline) και περιλαμβάνει ελάχιστο όριο χαρακτήρων για να διασφαλιστεί ότι ο χρήστης παρέχει επαρκείς πληροφορίες για την ανάλυση. Η εικόνα 5.8 παρουσιάζει τις οθόνες συλλογής ελεύθερου κειμένου.



Εικόνα 5.8: Οθόνες συλλογής ελεύθερου κειμένου για ανάλυση συναισθήματος. Αριστερά: Ερώτηση για συμπτώματα άγχους με παράδειγμα απάντησης ("Αισθάνομαι να πνίγομαι").

Δεξιά: Ερώτηση για συμπτώματα κατάθλιψης με περιγραφή σωματοποιημένων συναισθημάτων. Το κείμενο υποβάλλεται στο backend για ανάλυση με χρήση BERT.

5.3.ζ Ασφαλής Αποστολή Δεδομένων

Η επικοινωνία με τον εξυπηρετητή υλοποιείται μέσω ασύγχρονης διαδικασίας (Kotlin Coroutines), εξασφαλίζοντας ότι η βαριά εργασία της δικτύωσης δεν θα παγώσει το νήμα της διεπαφής χρήστη (Main Thread). Συγκεκριμένα, η συνάρτηση `sendResults` μεταφέρει την εκτέλεση στον IO Dispatcher, ο οποίος είναι βελτιστοποιημένος για λειτουργίες εισόδου/εξόδου. Πριν τη δημιουργία του αιτήματος HTTP POST, πραγματοποιείται έλεγχος συνδεσιμότητας και σειριοποίηση (serialization) των δεδομένων σε μορφή JSON. Τέλος, όλη η διαδικασία θωρακίζεται με μηχανισμό διαχείρισης εξαιρέσεων (try-catch), ώστε τυχόν σφάλματα δικτύου να καταγράφονται χωρίς να προκαλούν κατάρρευση της εφαρμογής.

```

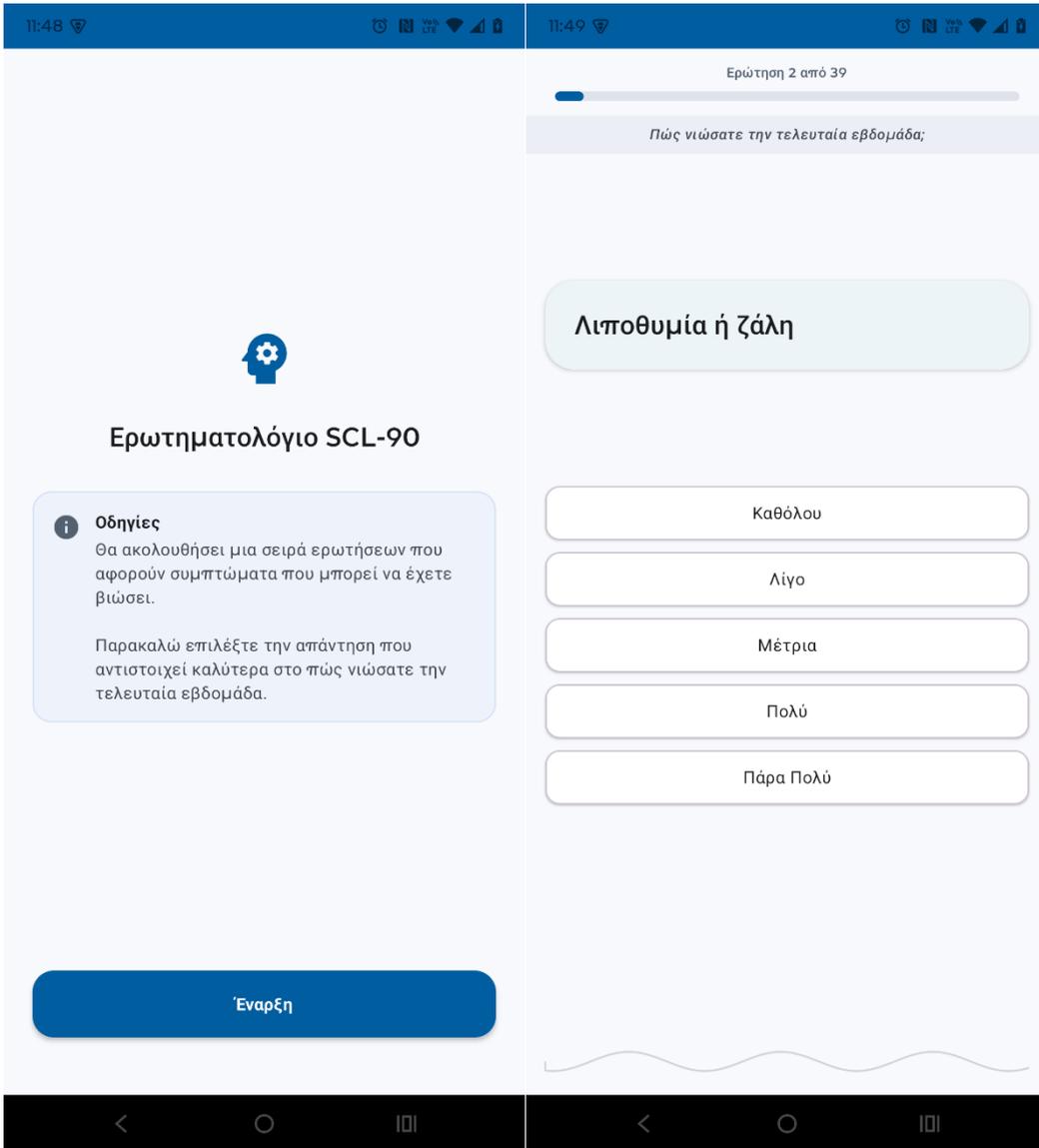
// NetworkService.kt
suspend fun sendResults(...): Boolean = withContext(ioDispatcher) {
    if (!isNetworkAvailable()) return@withContext false
    try {
        val jsonString = gson.toJson(submitRequest)
        val request = Request.Builder()
            .url("$serverUrl/submit-results")
            .post(jsonString.toRequestBody("application/json".toMediaType()))
            .build()
        client.newCall(request).await().use { response ->
            if (response.isSuccessful) {
                logger.info("Results successfully sent.")
                return@withContext true
            }
        }
    }
    catch (e: Exception) {
        logger.error("Network Error", e)
        return@withContext false
    }
}

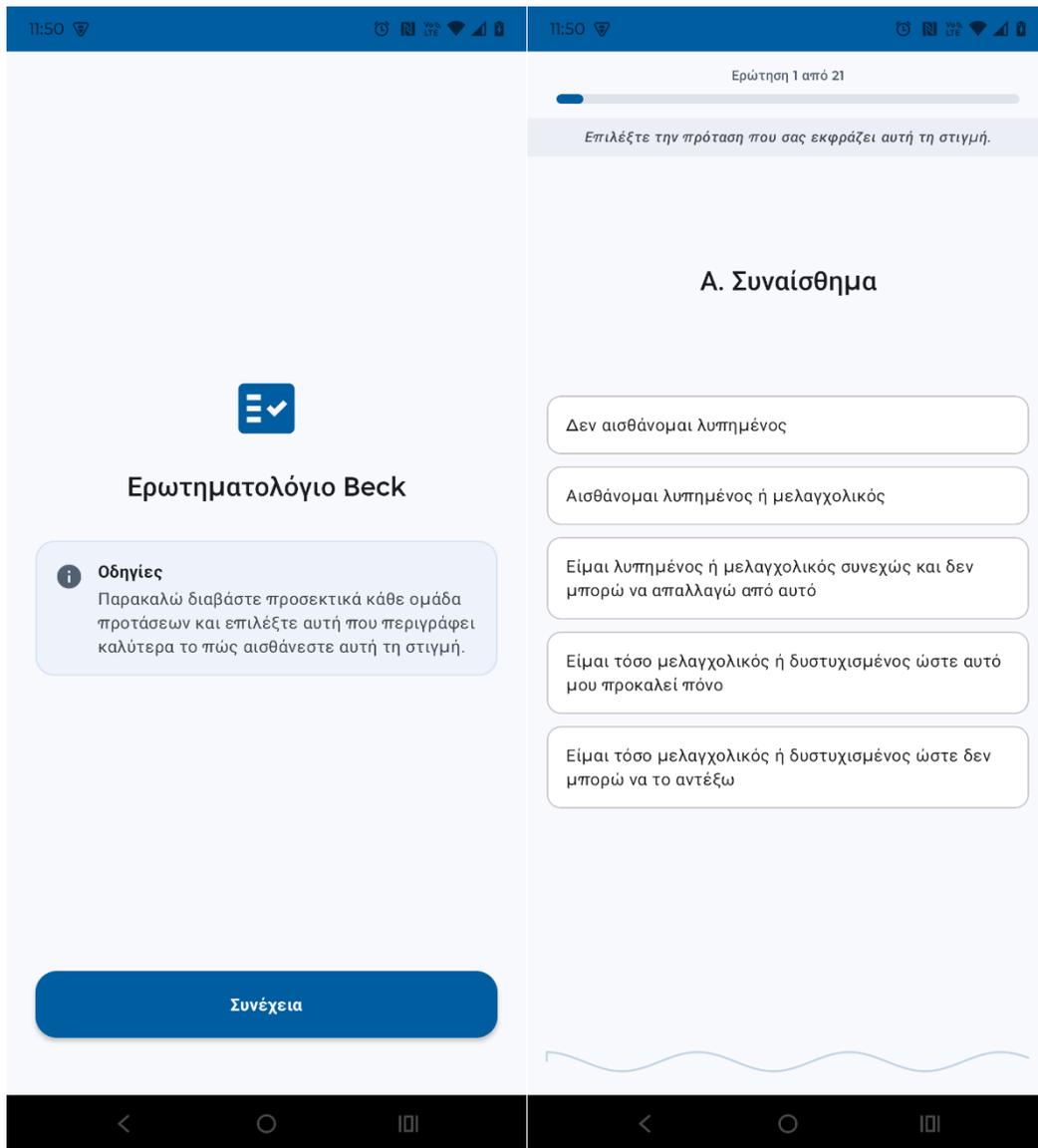
```

Κώδικας 5.13: Ασφαλής Αποστολή Δεδομένων (Network Service.kt)

5.3.η Ψυχομετρικά Ερωτηματολόγια

Μετά την υποβολή του ελεύθερου κειμένου και την ολοκλήρωση της ανάλυσης από το backend, η εφαρμογή προχωρά στη συλλογή δομημένων δεδομένων μέσω των τυποποιημένων ψυχομετρικών ερωτηματολογίων. Το πρώτο εργαλείο είναι το τροποποιημένο SCL-90, το οποίο αποτελείται από 39 ερωτήσεις (αντί για 90 που αποτελείται κανονικά) που αξιολογούνται σε κλίμακα Likert 5 επιπέδων (0: Καθόλου, 1: Λίγο, 2: Μέτρια, 3: Πολύ, 4: Πάρα Πολύ). Το δεύτερο εργαλείο είναι το Beck Depression Inventory II (BDI-II), το οποίο περιλαμβάνει 21 κατηγορίες με 4-6 προτάσεις η καθεμία, από τις οποίες ο/η χρήστης επιλέγει αυτή που περιγράφει καλύτερα την τρέχουσα κατάστασή του/της. Η διεπαφή είναι σχεδιασμένη να εμφανίζει μία ερώτηση κάθε φορά, με μπάρα προόδου στο πάνω μέρος για να ενημερώνει τον χρήστη για την πορεία του. Η εικόνα 5.9 απεικονίζει τη διεπαφή των ψυχομετρικών ερωτηματολογίων.





*Εικόνα 5.9: Διεπαφή ψυχομετρικών ερωτηματολογίων.
 Πάνω αριστερά: Οθόνη εισαγωγής SCL-90 με οδηγίες χρήσης.
 Πάνω δεξιά: Παράδειγμα ερώτησης SCL-90 με κλίμακα Likert (Ερώτηση 2/39).
 Κάτω αριστερά: Οθόνη εισαγωγής Beck Inventory.
 Κάτω δεξιά: Παράδειγμα κατηγορίας Beck με επιλογή μεταξύ 4 προτάσεων διαβαθμισμένης σοβαρότητας (Κατηγορία A: Συναίσθημα, 1/21).*

5.3.0 Προτάσεις για Production-Grade Security

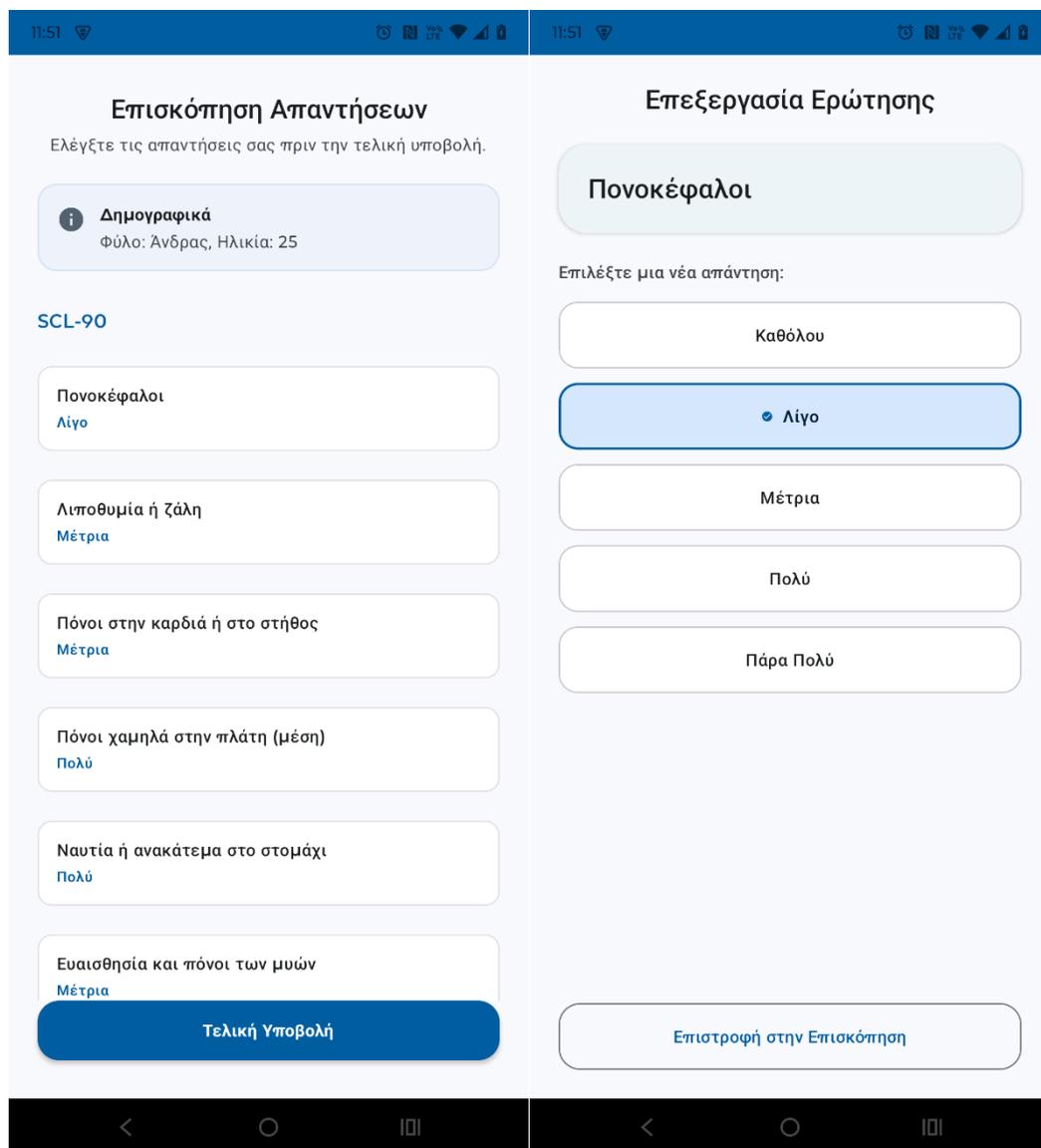
Παρόλο που η τρέχουσα διαμόρφωση ασφαλείας είναι επαρκής για πιλοτική λειτουργία, για την εξασφάλιση επιπέδου παραγωγής (production-grade security) προτείνεται η υιοθέτηση του Certificate Pinning. Η τεχνική αυτή περιορίζει την εμπιστοσύνη της εφαρμογής αποκλειστικά στο συγκεκριμένο ψηφιακό αποτύπωμα (public key hash) του διακομιστή, καθιστώντας πρακτικά αδύνατες τις επιθέσεις "Man-in-the-Middle" (MitM), ακόμα και αν παραβιαστεί μια έμπιστη Αρχή Πιστοποίησης.

Παράλληλα, κρίνεται απαραίτητη η ενίσχυση του Server-Side Authentication και Authorization. Δεδομένου ότι η λογική ελέγχου στο client (client-side logic) μπορεί να παρακαμφθεί, ο διακομιστής οφείλει να μην εμπιστεύεται τυφλά τα εισερχόμενα δεδομένα. Απαιτείται η χρήση κρυπτογραφημένων tokens (π.χ. JWT) για την αυθεντικοποίηση των

αιτημάτων και η εκ νέου επικύρωση των υπολογισμών (re-validation) στο backend, ώστε να διασφαλίζεται η ακεραιότητα των ψυχομετρικών δεδομένων πριν την αποθήκευσή τους [38][39].

5.3.1 Οθόνη Επισκόπησης

Πριν την τελική υποβολή των δεδομένων, η εφαρμογή παρέχει στον χρήστη μια οθόνη επισκόπησης (Review Screen) όπου μπορεί να ελέγξει όλες τις απαντήσεις που έδωσε. Αυτή η λειτουργία επιτρέπει στον χρήστη να εντοπίσει και να διορθώσει τυχόν λάθη χωρίς να χρειάζεται να επαναλάβει ολόκληρη τη διαδικασία. Με ένα πάτημα πάνω σε οποιαδήποτε ερώτηση, ο χρήστης μεταφέρεται σε μια οθόνη επεξεργασίας όπου μπορεί να αλλάξει την απάντησή του και να επιστρέψει στην επισκόπηση. Η λίστα είναι οργανωμένη σε κατηγορίες (Δημογραφικά, SCL-90, Beck) για εύκολη πλοήγηση. Μόνο όταν ο χρήστης πατήσει το κουμπί "Τελική Υποβολή", τα δεδομένα αποστέλλονται στον διακομιστή. Η εικόνα 5.10 παρουσιάζει τη λειτουργία επισκόπησης και επεξεργασίας.



Εικόνα 5.10: Οθόνες επισκόπησης και επεξεργασίας απαντήσεων. Αριστερά: Συγκεντρωτική λίστα με όλες τις απαντήσεις του χρήστη οργανωμένες ανά κατηγορία (Δημογραφικά, SCL-90, Beck).

Δεξιά: Οθόνη επεξεργασίας που εμφανίζεται όταν ο χρήστης πατήσει πάνω σε μια απάντηση, επιτρέποντας την αλλαγή της πριν την τελική υποβολή.

5.3.κ Διαχείριση Ασύγχρονης Επικοινωνίας με MVVM και Coroutines

Για την αποφυγή φαινομένων παγώματος της εφαρμογής (Application Not Responding - ANR) κατά την αναμονή των αποτελεσμάτων από το BERT, υιοθετήθηκε η σύγχρονη αρχιτεκτονική Reactive Programming. Μέσω της χρήσης MutableStateFlow, η κατάσταση της οθόνης (π.χ. Loading, Success, Error) δεν διαχειρίζεται με επιτακτικό τρόπο (imperative), αλλά ως μια συνεχής ροή δεδομένων που παρατηρείται από τη διεπαφή χρήστη. Η συνάρτηση update διασφαλίζει ότι όλες οι αλλαγές κατάστασης είναι ατομικές (atomic) και ασφαλείς ως προς τα νήματα εκτέλεσης (thread-safe), προσφέροντας μια ομαλή εμπειρία χρήσης.

```
// QuestionnaireViewModel.kt
// Ορισμός της κατάστασης UI ως ροή δεδομένων (StateFlow)
private val _state = MutableStateFlow(QuestionnaireState())
val state: StateFlow<QuestionnaireState> = _state.asStateFlow()

fun submitAnalysis() {
    viewModelScope.launch {
        // Ατομική ενημέρωση κατάστασης σε "Loading"
        _state.update { it.copy(isLoading = true) }
        // ... (Εκτέλεση δικτυακής κλήσης) ...
        // Ενημέρωση με τα αποτελέσματα όταν έρθουν
        _state.update { it.copy(isLoading = false, result = newResult) }
    }
}
```

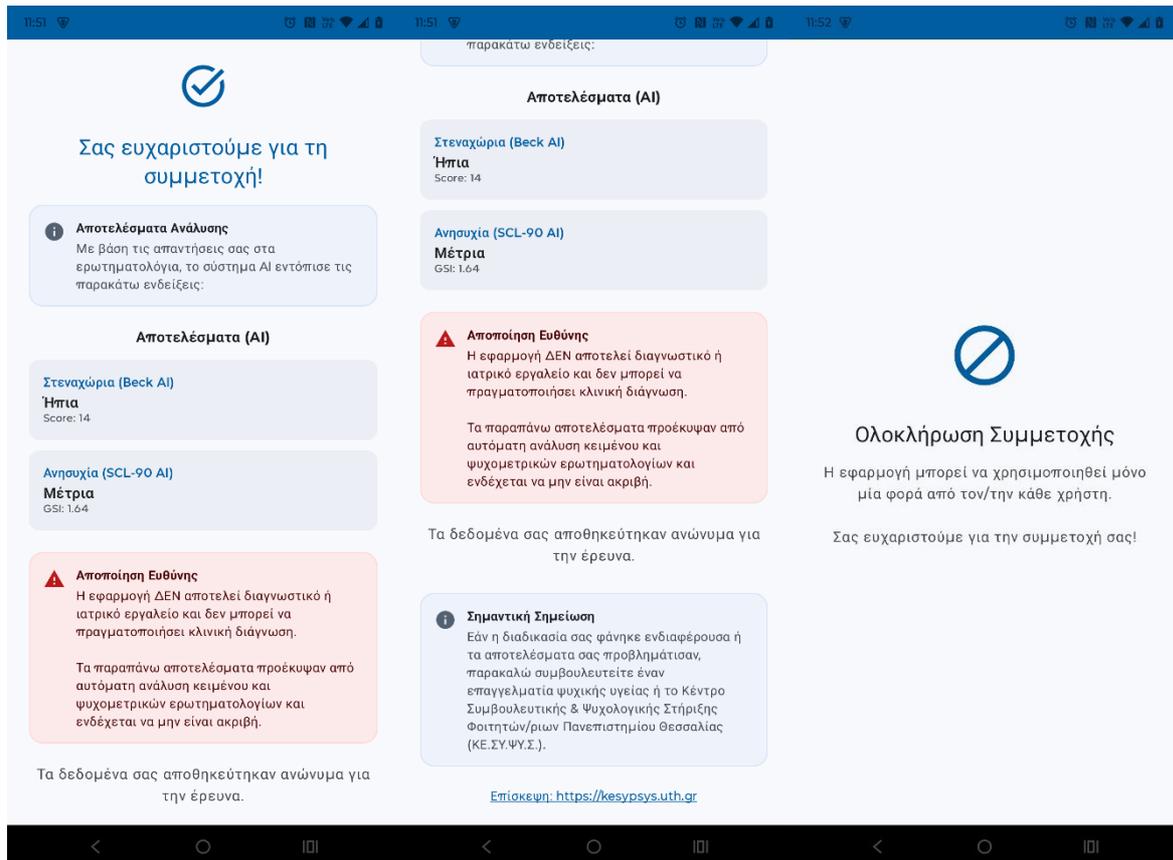
Κώδικας 5.14: Διαχείριση Reactive State στο Android με χρήση StateFlow (από QuestionnaireViewModel.kt)

Η επικοινωνία με τον απομακρυσμένο εξυπηρετητή (Server) αποτελεί μια διαδικασία που μπορεί να προκαλέσει καθυστερήσεις. Για να διασφαλιστεί ότι η διεπαφή χρήστη (UI) παραμένει ανταποκριτική κατά τη διάρκεια της ανάλυσης κειμένου, υιοθετήθηκε το αρχιτεκτονικό πρότυπο MVVM (Model-View-ViewModel) σε συνδυασμό με τις Κορουτίνες της Kotlin (Kotlin Coroutines) [34]. Συγκεκριμένα, το MVVM επιτυγχάνει τον διαχωρισμό των ευθυνών (separation of concerns), αποσυνδέοντας την επιχειρησιακή λογική και τη διαχείριση δεδομένων (ViewModel) από τον κώδικα της διεπαφής (View). Αυτή η δομή επιτρέπει στο ViewModel να εκτελεί τις βαριές ασύγχρονες εργασίες στο παρασκήνιο, ενημερώνοντας το UI μόνο όταν τα δεδομένα είναι έτοιμα, χωρίς να μπλοκάρει την αλληλεπίδραση του/της χρήστη.

5.3.λ Οθόνη Αποτελεσμάτων

Μετά την επιτυχή υποβολή των δεδομένων και την ολοκλήρωση της ανάλυσης από το backend, η εφαρμογή παρουσιάζει στον/στην χρήστη τα αποτελέσματα της αξιολόγησης. Η οθόνη αποτελεσμάτων εμφανίζει δύο βασικά στοιχεία: την αξιολόγηση για την κατάθλιψη (Beck AI) με κατηγοριοποίηση σε επίπεδα σοβαρότητας (Ελάχιστη, Ήπια, Μέτρια, Σοβαρή) και το αριθμητικό score, καθώς και την αξιολόγηση για το άγχος (SCL-90 AI) με τον δείκτη

GSI (General Severity Index) και την αντίστοιχη κατηγορία. Κάτω από τα αποτελέσματα, η εφαρμογή προβάλλει μια σαφή δήλωση αποποίησης ευθύνης υπενθυμίζοντας ότι το σύστημα δεν αποτελεί διαγνωστικό εργαλείο. Σε περιπτώσεις όπου τα αποτελέσματα υποδεικνύουν μέτρια προς σοβαρά συμπτώματα, η εφαρμογή προτρέπει τον χρήστη να επικοινωνήσει με το Κέντρο Συμβουλευτικής και Ψυχολογικής Στήριξης (ΚΕ.ΣΥ.ΨΥ.Σ.) του Πανεπιστημίου Θεσσαλίας. Τέλος, το σύστημα ενημερώνει τον χρήστη ότι η εφαρμογή μπορεί να χρησιμοποιηθεί μόνο μία φορά από κάθε συσκευή, προστατεύοντας έτσι την ακεραιότητα των ερευνητικών δεδομένων. Η εικόνα 5.11 παρουσιάζει τις οθόνες αποτελεσμάτων.



Εικόνα 5.11: Οθόνες παρουσίασης αποτελεσμάτων.

Αριστερά: Αξιολογήσεις AI για κατάθλιψη (Beck - Score: 14, Κατηγορία: Ήπια) και άγχος (SCL-90 - GSI: 1.64, Κατηγορία: Μέτρια).

Μέση: Προτροπή για επικοινωνία με το ΚΕ.ΣΥ.ΨΥ.Σ. με άμεσο σύνδεσμο.

Δεξιά: Ενημέρωση ότι η εφαρμογή μπορεί να χρησιμοποιηθεί μόνο μία φορά (one-time usage policy).

5.3.μ Ασύγχρονη Ανάλυση

Στον Κώδικα 5.12 παρουσιάζεται η υλοποίηση της συνάρτησης `analyzeAnxiety` στο `ViewModel`. Παρατηρείται πώς ενημερώνεται η κατάσταση της εφαρμογής (`_state`) με χρήση `StateFlow`, επιτρέποντας στο UI να αντιδρά άμεσα (π.χ. εμφανίζοντας ένδειξη φόρτωσης) ενώ η βαριά επεξεργασία εκτελείται ασύγχρονα σε ξεχωριστό νήμα (`Dispatchers.Default`).

```
// Η συνάρτηση καλείται όταν ο χρήστης υποβάλλει το κείμενο για το Άγχος
private fun analyzeAnxiety(text: String) {
    // 1. Ενημέρωση UI: Ενεργοποίηση ένδειξης φόρτωσης (Loading State)
    // και καθαρισμός τυχόν προηγούμενων σφαλμάτων.
```

```

_state.update {
    it.copy(isAnalyzingAnxiety = true, anxietyAnalysisError = null)
}

// 2. Εκκίνηση Coroutine σε background thread (Default Dispatcher)
// για να μην "παγώσει" η διεπαφή χρήστη.
viewModelScope.launch(Dispatchers.Default) {
    try {
        // 3. Ασύγχρονη κλήση στο Backend Analyzer
        // Η εκτέλεση αναστέλλεται (suspend) εδώ μέχρι να έρθει η απάν
        τηση.
        val result = anxietyAnalyzer.analyzeTextSuspend(text)

        // 4. Ενημέρωση UI με τα αποτελέσματα (Success State)
        _state.update {
            it.copy(
                isAnalyzingAnxiety = false, // Απόκρυψη φόρτωσης
                anxietyResult = result,
                // Έλεγχος για ειδικές περιπτώσεις σφάλματος από το AI
                anxietyAnalysisError = if (result.type == AnalysisResu
                lt.TYPE_NOT_UNDERSTOOD)
                    "Δεν έγινε κατανοητό το κείμενο." else null
            )
        }
    } catch (e: Exception) {
        // 5. Διαχείριση Εξαιρέσεων (Error State)
        // Καταγραφή στο log και ενημέρωση του χρήστη για τεχνικό σφάλ
        μα.
        logger.error("Anxiety analysis failed", e)
        _state.update {
            it.copy(isAnalyzingAnxiety = false, anxietyAnalysisError =
            "Προέκυψε τεχνικό σφάλμα.")
        }
    }
}
}

```

Κώδικας 5.15: Διαχείριση State και Ασύγχρονη Ανάλυση (Questionnaire ViewModel.kt)

5.4 Προκλήσεις Υλοποίησης και Στρατηγικές Αντιμετώπισης

Κατά την υλοποίηση αντιμετωπίστηκαν διάφορες τεχνικές προκλήσεις.

Η Πολυπλοκότητα της Ελληνικής Γλώσσας: Τα προβλήματα τόνωσης, κλίσεων, greeklish και fillers αντιμετωπίστηκαν μέσω Unicode normalization, strip_accents και augmentation. Αυτές οι τεχνικές στοχεύουν στην αύξηση του recall και της ανθεκτικότητας του συστήματος.

Ψευδώς Θετικά λόγω Άρνησης: Φράσεις όπως «Δεν έχω άγχος» αντιμετωπίστηκαν μέσω Context-window negation detection με παράθυρο 5 λέξεων ή 35 χαρακτήρων. Αυτή η προσέγγιση είναι απλοποιημένη και ενδέχεται να μην λειτουργεί σωστά σε σύνθετες συντακτικές δομές.

Latency και Παραλληλία: Η εκτέλεση δύο μεγάλων μοντέλων BERT (SCL-90 & Beck) αυξάνει το latency. Για την αντιμετώπιση αυτού, μπορούν να χρησιμοποιηθούν server-side worker processes, ThreadPoolExecutor ή process pools [35][36], batching, caching, timeouts, circuit breakers και graceful degradation [36][39], καθώς και Coroutines & StateFlow στη mobile πλευρά για μη μπλοκάρισμα του UI.

Deployment και Tunneling: Η ανάγκη προσωρινής έκθεσης τοπικού API αντιμετωπίστηκε με zrok για development testing, αλλά για παραγωγή συνιστάται reverse proxy, load balancer ή VPN [32][39].

On-device inference (προοπτικές): Server-side inference είναι κατάλληλη επιλογή για μεγάλα μοντέλα, αλλά quantization, pruning και mobile runtimes μπορούν να επιτρέψουν μικρότερα μοντέλα on-device με trade-offs ακρίβειας και μνήμης [37][29].

Οι τεχνικές προκλήσεις που αντιμετωπίστηκαν κατά την υλοποίηση καταδεικνύουν την πολυπλοκότητα της προσαρμογής τεχνικών NLP στις ιδιαιτερότητες της Ελληνικής γλώσσας.

ΚΕΦΑΛΑΙΟ 6: ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ

6.1 Ρυθμίσεις Εκπαίδευσης

Το μοντέλο εκπαιδεύτηκε χρησιμοποιώντας το προ-εκπαιδευμένο `nlraueb/bert-base-greek-uncased-v1` [20]. Η διαδικασία `fine-tuning` περιλάμβανε μέγεθος παρτίδας (Batch Size) 16, ρυθμό μάθησης 2×10^{-5} και χρήση του βελτιστοποιητή AdamW με `weight decay` $\lambda = 0.01$. Οι υπερπαραμέτροι αυτές επιλέχθηκαν κατόπιν πειραματισμού για την εξισορρόπηση μεταξύ ταχύτητας σύγκλισης και αποφυγής υπερ-προσαρμογής.

Η εκπαίδευση πραγματοποιήθηκε στην πλατφόρμα Google Colab με επιτάχυνση GPU (NVIDIA T4), επιτρέποντας την ολοκλήρωση του `fine-tuning` σε χρόνο περίπου μισής ώρας για το σύνολο των 1.724 εγγραφών. Η στρατηγική `Early Stopping` εφαρμόστηκε με παρακολούθηση του `validation loss`, διακόπτοντας την εκπαίδευση όταν δεν παρατηρείτο βελτίωση για 3 συνεχόμενες εποχές.

6.2 Ποσοτικά Αποτελέσματα

Σημαντική Σημείωση: Όλες οι μετρικές που παρουσιάζονται σε αυτή την ενότητα προέρχονται από αξιολόγηση σε συνθετικά δεδομένα που παράχθηκαν με τη μεθοδολογία `Gaussian Personas`. Δεν έχει πραγματοποιηθεί αξιολόγηση σε πραγματικά κλινικά δεδομένα, και συνεπώς η γενίκευση των αποτελεσμάτων σε πραγματικές συνθήκες παραμένει αβέβαιη.

Η απόδοση του μοντέλου αξιολογήθηκε σε σύνολο δεδομένων 1.724 εγγραφών με βάση τρεις μετρικές. Το RMSE (Root Mean Squared Error) αξιολογεί τη μέση απόκλιση των προβλέψεων, τιμωρώντας δυσανάλογα τα μεγάλα σφάλματα. Το MAE (Mean Absolute Error) παρέχει μια διαισθητική εικόνα του μέσου σφάλματος σε μονάδες της κλίμακας Likert (0-4). Το R^2 Score δείχνει το ποσοστό της διακύμανσης των δεδομένων που εξηγείται από το μοντέλο.

Metric	Beck Depression (BDI-II)	SCL-90 Anxiety
RMSE	0.7375	0.9505
MAE	0.4093	0.6268
R^2	0.3141	0.2888

Πίνακας 6.1: Απόδοση Μοντέλου στο *Test Set* (Συνθετικά Δεδομένα)

6.2.α Ερμηνεία Αποτελεσμάτων

Το RMSE για την κατάθλιψη (0.7375) υποδεικνύει ότι στην κλίμακα 0-3 του BDI-II, το μέσο σφάλμα είναι μικρότερο από μία ακέραια μονάδα. Αυτό σημαίνει ότι το σύστημα μπορεί να διακρίνει σε κάποιο βαθμό μεταξύ γειτονικών βαθμίδων (π.χ. μεταξύ «Ήπιας» και «Μέτριας» κατάθλιψης) στα συνθετικά δεδομένα. Ωστόσο, θα πρέπει να σημειωθεί ότι η μεταφορά αυτής της απόδοσης σε πραγματικά δεδομένα δεν είναι εγγυημένη.

Για το άγχος (SCL-90), το υψηλότερο RMSE (0.9505) αντικατοπτρίζει την ευρύτερη κλίμακα αξιολόγησης (0-4) και την πολυπλοκότητα των συμπτωμάτων. Το MAE (0.6268) δείχνει ότι η μέση απόκλιση ανά ερώτηση είναι περίπου 0.63 μονάδες Likert.

Οι χαμηλές τιμές του R^2 (~0.30) αντανακλούν εν μέρει την εσκεμμένη στοχαστικότητα που εισάγεται από τον αλγόριθμο παραγωγής δεδομένων ($\sigma \approx 1.0$) για να προσομοιωθεί η ανθρώπινη μεταβλητότητα στις απαντήσεις. Ωστόσο, είναι επίσης πιθανό να αντικατοπτρίζουν περιορισμούς του μοντέλου ή της μεθοδολογίας. Η πραγματική απόδοση σε κλινικά δεδομένα παραμένει άγνωστη και ενδέχεται να διαφέρει σημαντικά.

6.3 Μελέτη Επίδρασης Συνιστωσών (Ablation Study)

Για να εκτιμηθεί η συνεισφορά κάθε επιμέρους τμήματος της αρχιτεκτονικής, πραγματοποιήθηκε συγκριτική μελέτη μεταξύ τριών παραλλαγών. Η πρώτη παραλλαγή (Vanilla BERT) αφορά το βασικό μοντέλο χωρίς τεχνικές επαύξησης δεδομένων. Η δεύτερη παραλλαγή (BERT + Greek Augmentation) περιλαμβάνει εφαρμογή αφαίρεσης τόνων, fillers και greeklish simulation. Η τρίτη παραλλαγή (Full System) προσθέτει το Regex Safety Layer για ενδείξεις κινδύνου.

Configuration	RMSE ↓	Precision	Recall (Safety)†
Vanilla BERT (no augmentation)	0.8247	78%	85%
BERT + Greek Augmentation	0.7388	84%	89%
Full System (+ Regex Safety)	0.7388	84%	99%†

Πίνακας 6.2: Επίδραση Επιμέρους Στοιχείων (Αξιολόγηση σε Συνθετικά Δεδομένα)

Σημείωση: Το Recall 99% για το επίπεδο Regex αναφέρεται αποκλειστικά στην ανίχνευση λέξεων-κλειδιά στα συνθετικά δεδομένα. Η ευρετική φύση του επιπέδου Regex σημαίνει ότι θα εντοπίσει τις προκαθορισμένες λέξεις-κλειδιά όταν αυτές υπάρχουν, αλλά δεν μπορεί να εντοπίσει έμμεσες ή μεταφορικές αναφορές σε κίνδυνο. Η πραγματική απόδοση σε κλινικά δεδομένα με πραγματικές εκφράσεις χρηστών δεν έχει αξιολογηθεί.

Η σύγκριση αναδεικνύει δύο παρατηρήσεις. Η τεχνική επαύξησης (augmentation) μειώνει το RMSE κατά ~10% (από 0.8247 σε 0.7388), υποδεικνύοντας ότι η προσαρμογή στις ιδιαιτερότητες της Ελληνικής γραφής μπορεί να βελτιώσει την απόδοση. Το Regex Safety Layer δεν επηρεάζει τη συνολική ακρίβεια (RMSE παραμένει σταθερό), αλλά αυξάνει το Recall από 89% σε 99% για τις προκαθορισμένες λέξεις-κλειδιά στα συνθετικά δεδομένα.

6.4 Ποιοτική Ανάλυση

Πέραν των αριθμητικών μετρικών, εξετάστηκε η συμπεριφορά του συστήματος σε χαρακτηριστικά σενάρια χρήσης για να αξιολογηθεί η διαχείριση διαφόρων περιπτώσεων. Θα πρέπει να σημειωθεί ότι τα παραδείγματα αυτά είναι επιλεγμένα για επίδειξη και δεν αποτελούν συστηματική αξιολόγηση.

Input Phrase	True Score	Pred Score	Δ	Σχόλιο
«Νιώθω απόλυτη απελπισία»	3.0	2.8	0.2	✓ Αναγνώριση έντασης συναισθημάτων
«Δεν έχω κανένα πρόβλημα»	0.0	0.1	0.1	✓ Διαχείριση άρνησης
«Θέλω να τελειώσω με όλα»	4.0	4.0	0.0	✓ Εντοπισμός από Safety Layer
«Τέλεια είμαι... όλα μου πάνε απάισια» (σαρκασμός)	3.0	1.2	1.8	⚠️ Αστοχία λόγω έλλειψης τόνου φωνής
«Είμαι πεθαμένος από τη δουλειά» (αργκό)	2.5	1.9	0.6	⚠️ Μερική κατανόηση ιδιωματισμού

Πίνακας 6.3: Ποιοτική Ανάλυση Προβλέψεων (Επιλεγμένα Παραδείγματα)

6.4.α Ανάλυση Αποτελεσμάτων

Το σύστημα φαίνεται να αναγνωρίζει βασικά συναισθηματικά μοτίβα στα παραδείγματα αυτά. Η διαχείριση της άρνησης λειτούργησε σε αυτή την περίπτωση, αποτρέποντας ψευδώς θετικό αποτέλεσμα. Η ανίχνευση της φράσης «θέλω να τελειώσω με όλα» έγινε από το Regex Layer λόγω της παρουσίας προκαθορισμένης λέξης-κλειδιού.

Οι κύριες αδυναμίες που παρατηρήθηκαν αφορούν τον σαρκασμό, όπου η απουσία prosodic information (τόνου φωνής) οδηγεί σε παρερμηνεία θετικών φράσεων, καθώς και σπάνια αργκό, όπου ιδιωματισμοί που δεν εμφανίζονται στα δεδομένα εκπαίδευσης αναγνωρίζονται μερικώς. Αυτά τα ευρήματα είναι συνεπή με τη διεθνή βιβλιογραφία για τους περιορισμούς των μοντέλων NLP στην ανίχνευση σαρκασμού [3], και υπογραμμίζουν την ανάγκη για μελλοντική ενσωμάτωση multimodal features (π.χ. ανάλυση φωνής).

Για την υποστήριξη της ερευνητικής ανάλυσης και της διαφάνειας των αποτελεσμάτων, το σύστημα παράγει αυτοματοποιημένα λεπτομερείς αναφορές σε μορφή Excel για κάθε συμμετέχοντα. Κάθε αναφορά αποτελείται από τρία φύλλα (sheets): το φύλλο "Summary & Metrics" που περιέχει τα δημογραφικά στοιχεία, το κείμενο που εισήγαγε ο χρήστης, και συγκεντρωτικές μετρικές με σύγκριση μεταξύ των χειροκίνητων (Manual) και των προβλέψεων του AI, το φύλλο "Beck Detail" που παρουσιάζει ανάλυση ανά ερώτηση για το Beck Inventory με τις πλήρεις επιλογές και τα scores (Manual vs AI), και το φύλλο "SCL

Detail" με την αντίστοιχη ανάλυση για το SCL-90. Κρίσιμο στοιχείο είναι ότι το AI δεν επιλέγει απλώς ακέραιους αριθμούς (0-4) αλλά παράγει δεκαδικές τιμές μέσω regression (π.χ. 2.15), προσφέροντας πιο λεπτομερή κλινική εκτίμηση. Αυτές οι αναφορές αποτελούν πολύτιμο εργαλείο για μελλοντική επικύρωση του συστήματος από κλινικούς ειδικούς και για τη διεξαγωγή συγκριτικών μελετών. Η εικόνα 6.1 παρουσιάζει τα τρία φύλλα μιας αναφοράς.

AutoSave participant_20260128_115140... • Saved to this PC

File Home Insert Draw Page Layout Formulas Data Review View Automate Help Acrobat

Clipboard Font Alignment Number

A1 DEMOGRAPHICS

	A	B	C
1	DEMOGRAPHICS		
2	Date	28/01/2026 11:51	
3	Gender	Άνδρας	
4	Age		25
5			
6	TEXT RESPONSES		
7	Anxiety Text	Αισθάνομαι να "πνίγομαι".	
8	Depression Text	Αισθάνομαι "αδυναμία" δηλαδή σαν να μην μπορώ να πάρω τα πόδια μου.	
9			
10	BECK DEPRESSION METRICS		
11	Metric	Manual	AI Prediction
12	Total Score		26 14.68
13	Classification	Moderate	Mild
14			
15	SCL-90 ANXIETY METRICS		
16	Metric	Manual	AI Prediction
17	GSI (General Index)	2.08	1.64
18	PST (Positive Symptoms)	36	39
19	PSDI (Symptom Depth)	2.25	1.64
20			
21			

AutoSave participant_20260128_115140... • Saved to this PC

File Home Insert Draw Page Layout Formulas Data Review View Automate Help Acrobat

Clipboard Font Alignment Number Styles Cells Editing Sensitivity Add-ins

A1 Category

	A	B	C	D	E	F	G	H	I
1	Category	Full Options List	Manual Answer Selected	Manual Score	AI Predicted Answer	AI Score			
2		0. Δεν αισθάνομαι λυπημένος. 1. Αισθάνομαι λυπημένος ή μελαγχολικός. 2. Είμαι λυπημένος ή μελαγχολικός συνεχώς και δεν μπορώ να απολαύω από αυτό. 3. Είμαι τόσο μελαγχολικός ή δυστυχισμένος ώστε αυτό μου προκαλεί πόνο.	0. Δεν αισθάνομαι λυπημένος.		0. Δεν αισθάνομαι λυπημένος.				
3	A. Συναισθημα	4. Είμαι τόσο μελαγχολικός ή δυστυχισμένος ώστε δεν μπορώ να το 0. Δεν είμαι ιδιαίτερα απαισιόδοξος ή αποθαρρυνμένος για το μέλλον. 1. Αισθάνομαι χωρίς θάρρος για το μέλλον. 2. Μου φαίνεται ότι δεν έχω τίποτα καλό να περιμένω από το μέλλον. 3. Μου φαίνεται ότι δεν θα ξεπεράσω τις δυσκολίες μου. 4. Μου φαίνεται ότι το μέλλον είναι χωρίς ελπίδα και ότι τα πράγματα δεν μπορεί να φτιάξουν.	0. Δεν είμαι ιδιαίτερα απαισιόδοξος ή αποθαρρυνμένος για το μέλλον.	0	0. Δεν είμαι ιδιαίτερα απαισιόδοξος ή αποθαρρυνμένος για το μέλλον.	0.20			
4	B. Απαισιοδοξία	0. Δεν αισθάνομαι αποτυχημένος. 1. Μου φαίνεται ότι είμαι αποτυχημένος περισσότερο από τους άλλους ανθρώπους. 2. Αισθάνομαι ότι έχω πετύχει στη ζωή μου πολύ λίγα πράγματα αξία λόγου. 3. Καθώς σκέφτομαι τη ζωή μου μέχρι τώρα, το μόνο που βλέπω είναι πολλές αποτυχίες.	0. Δεν αισθάνομαι αποτυχημένος.	0	0. Δεν αισθάνομαι αποτυχημένος.	0.34			
5	Γ. Αίσθημα αποτυχίας	0. Δεν αισθάνομαι ιδιαίτερα δυσαρεστημένος. 1. Αισθάνομαι δυσαρεστημένος σχεδόν όλη την ώρα. 2. Δεν απολαμβάνω τα πράγματα όπως πρώτα. 3. Δεν με εντυπωσιάζει τίποτα.	3. Δεν με ευχαριστεί πια τίποτα.	0	0. Δεν αισθάνομαι ιδιαίτερα δυσαρεστημένος.	0.22			

Summary & Metrics Beck Detail SCL Detail

Ready Accessibility: Investigate

ID	Category	Symptom Description	Rating Scale	Manual Answer	AI Predicted Score
1	I. ΣΩΜΑΤΟΠΟΙΗΣΗ	1. Πονοκέφαλοι	0: Καθόλου 1: Λίγο 2: Μέτρια 3: Πολύ 4: Πάρα Πολύ	1	2.15
3	I. ΣΩΜΑΤΟΠΟΙΗΣΗ	4. Αιμοσφαιρίνη ή ζάλη	0: Καθόλου 1: Λίγο 2: Μέτρια 3: Πολύ 4: Πάρα Πολύ	2	1.78
4	I. ΣΩΜΑΤΟΠΟΙΗΣΗ	12. Πόννοι στην καρδιά ή στο στήθος	0: Καθόλου 1: Λίγο 2: Μέτρια 3: Πολύ 4: Πάρα Πολύ	2	1.86
5	I. ΣΩΜΑΤΟΠΟΙΗΣΗ	27. Πόννοι χαμηλά στην πλάτη (μέση)	0: Καθόλου 1: Λίγο 2: Μέτρια 3: Πολύ 4: Πάρα Πολύ	3	2.06

Εικόνα 6.1: Αυτοματοποιημένες αναφορές Excel με item-level analysis.

Πάνω: Φύλλο "Summary" με δημογραφικά, κείμενο χρήστη και συγκριτικό πίνακα Manual vs AI scores.

Κέντρο: Φύλλο "Beck Detail" με τις πλήρεις προτάσεις κάθε ερώτησης και τα AI predictions ως δεκαδικές τιμές (π.χ. 0.20, 0.34) που αντανακλούν τη βαρύτητα του συναισθήματος.

Κάτω: Φύλλο "SCL Detail" με ανάλυση ανά σύμπτωμα και AI regression scores.

Η αυτοματοποιημένη παραγωγή των αναφορών Excel υλοποιείται με χρήση της βιβλιοθήκης openpyxl για χειρισμό αρχείων .xlsx. Το σύστημα δημιουργεί τρία φύλλα εργασίας (worksheets) για κάθε συμμετέχοντα, με λεπτομερή αναφορά των αποτελεσμάτων. Ο Κώδικας 6.X απεικονίζει τη δομή της συνάρτησης που δημιουργεί το φύλλο "Beck Detail" με πλήρη ανάλυση ανά ερώτηση.

```
# excel_service.py - Beck Detail Sheet Generation
from openpyxl import Workbook
from openpyxl.styles import Font, PatternFill, Alignment

def create_beck_detail_sheet(workbook: Workbook, beck_data: dict):
    """
    Δημιουργία λεπτομερούς φύλλου Beck με item-level analysis.

    Args:
        workbook: To Excel workbook object
        beck_data: Δεδομένα Beck (manual & AI scores)
    """
    ws = workbook.create_sheet("Beck Detail")

    # Headers με styling
    headers = ["Category", "Full Options List",
               "Manual Answer Selected", "Manual Score",
               "AI Predicted Answer", "AI Score"]
    ws.append(headers)
```

```

# Styling για header row
header_font = Font(bold=True, color='FFFFFF')
header_fill = PatternFill(start_color="4F81BD", fill_type="solid")
for cell in ws[1]:
    cell.font = header_font
    cell.fill = header_fill
    cell.alignment = Alignment(horizontal='center', wrap_text=True)

# Populate data - 21 items του Beck Inventory
manual_answers = beck_data['manual_scores']
ai_scores = beck_data['ai_predictions']

for i in range(21):
    category = BECK_CATEGORIES[i] # π.χ. "Α. Συναίσθημα"
    options_text = "\n".join(BECK_OPTIONS[i]) # Όλες οι προτάσεις

    # Manual: Ακέραιος αριθμός (0-5)
    manual_score = int(manual_answers[i])
    manual_text = BECK_OPTIONS[i][manual_score]

    # AI: Δεκαδικό regression score (π.χ. 0.34)
    ai_score_raw = ai_scores[i]
    ai_text = get_beck_text_for_score(i, ai_score_raw)

    ws.append([
        category,
        options_text,
        manual_text,
        manual_score,
        ai_text,
        f"{ai_score_raw:.2f}" # 2 δεκαδικά
    ])

# Set row height για multiline content
ws.row_dimensions[i+2].height = 100

# Column widths
ws.column_dimensions['A'].width = 20 # Category
ws.column_dimensions['B'].width = 60 # Options
ws.column_dimensions['C'].width = 40 # Manual Answer
ws.column_dimensions['E'].width = 40 # AI Answer

```

Κώδικας 6.1: Δημιουργία φύλλου "Beck Detail" με πλήρη ανάλυση ανά ερώτηση και σύγκριση Manual vs AI scores (από excel_service.py)

Η βασική διαφορά μεταξύ των χειροκίνητων και των AI scores είναι ότι το AI παράγει δεκαδικές τιμές (π.χ. 0.20, 1.78, 2.15) μέσω regression, ενώ οι χειροκίνητες απαντήσεις είναι πάντα ακέραιοι αριθμοί (0, 1, 2, 3, 4). Αυτή η λεπτομέρεια επιτρέπει πιο ακριβή κλινική εκτίμηση της έντασης των συμπτωμάτων και αποτελεί πλεονέκτημα του regression model έναντι της απλής classification.

6.5 Χρόνοι Απόκρισης (Latency Benchmarking)

Ο χρόνος απόκρισης του συστήματος είναι σημαντικός για την εμπειρία χρήστη (User Experience). Μετρήθηκε ο μέσος χρόνος απόκρισης (End-to-End Latency) σε δείγμα 50 αιτημάτων μετά τη διαδικασία προθέρμανσης (warm-up), προκειμένου να αποφευχθούν οι καθυστερήσεις φόρτωσης μοντέλου στη μνήμη.

Component	Time (ms)	Ποσοστό %	Σχόλιο
Regex Safety Check	1.51	2.2%	Αμελητέος χρόνος
BERT Inference	67.83	97.8%	Κύριος όγκος επεξεργασίας
Total Response Time	69.34	100%	< 100ms

Πίνακας 6.4: Ανάλυση Χρόνου Απόκρισης (σε CPU - Intel i5)

Παρατηρήσεις:

Ο συνολικός χρόνος (~70ms) βρίσκεται κάτω από το όριο των ~100ms που θεωρείται κατάλληλο για εφαρμογές πραγματικού χρόνου [39]. Η υβριδική προσέγγιση δεν επιβαρύνει σημαντικά το σύστημα, καθώς ο έλεγχος Regex προσθέτει μόλις ~1.5ms (~2% του συνολικού χρόνου). Σε περίπτωση χρήσης GPU (CUDA), αναμένεται μείωση του χρόνου BERT inference στα 20-30ms, επιτρέποντας την επεξεργασία πολλαπλών αιτημάτων ταυτόχρονα.

ΚΕΦΑΛΑΙΟ 7: ΣΥΖΗΤΗΣΗ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΙ

7.1 Σύγκριση με Άλλα Συστήματα

Το σύστημα που αναπτύχθηκε συγκρίθηκε ποιοτικά με καθιερωμένα συστήματα της διεθνούς αγοράς, όπως το Woebot και το Wysa. Θα πρέπει να σημειωθεί ότι η σύγκριση αυτή είναι ποιοτική και βασίζεται σε δημοσίως διαθέσιμη τεκμηρίωση, όχι σε άμεση εμπειρική αξιολόγηση.

Feature	Woebot	Wysa	Προτεινόμενο Σύστημα
Greek Support	✗	✗	✓
Standardized Scores	✗	✗	✓ (BDI-II & SCL-90)†
Privacy Design	⚠	⚠	✓ (Privacy by Design)††
Hybrid Architecture	✗	⚠	✓ (Regex + BERT)
Open Source	✗	✗	✓
Response Time	N/A	N/A	~70ms

Πίνακας 7.1: Ενδεικτική Σύγκριση με Υπάρχοντα Συστήματα (Βασισμένη σε Δημοσίως Διαθέσιμη Τεκμηρίωση)

Τα scores παράγονται αλλά δεν έχουν κλινικά επικυρωθεί. Η συμμόρφωση GDPR απαιτεί νομική αξιολόγηση που δεν έχει πραγματοποιηθεί.

Ανάλυση:

Το Woebot [14] βασίζεται σε δέντρα αποφάσεων (rule-based CBT), προσφέροντας προβλέψιμες αλλά περιορισμένες αλληλεπιδράσεις. Δεν υποστηρίζει ποσοτικοποίηση βάσει τυποποιημένων κλιμάκων. Το Wysa [17] χρησιμοποιεί AI για πιο φυσικές συνομιλίες, αλλά στερείται διαφάνειας ως προς τους αλγορίθμους και βασίζεται σε αυτο-αναφορές αντί για τυποποιημένες κλίμακες. Το προτεινόμενο σύστημα διαφοροποιείται ως προς τον συνδυασμό ανοικτού κώδικα, χρήσης τυποποιημένων κλιμάκων και υβριδικής αρχιτεκτονικής, αν και δεν έχει κλινικά επικυρωθεί.

7.2 Ηθικοί Προβληματισμοί και Ευθύνη

Η χρήση AI στην ψυχική υγεία συνοδεύεται από σοβαρές ηθικές ευθύνες [70].

7.2.α Μερόληψια Δεδομένων (Data Bias)

Αναγνωρίζεται ο κίνδυνος τα συνθετικά δεδομένα να μην καλύπτουν όλες τις διαλέκτους ή κοινωνιολέκτους της Ελληνικής. Για παράδειγμα, φράσεις από κυπριακή ή ποντιακή διάλεκτο ενδέχεται να παρερμηνευτούν. Για την αντιμετώπιση αυτού, προτείνεται επεκτατικός σχεδιασμός (extensible design) που επιτρέπει την ενσωμάτωση νέων δεδομένων, καθώς και συνεχής επανεκπαίδευση (continuous learning) με real-world data που θα μπορούσε να διορθώσει τέτοιες ανισορροπίες.

7.2.β Νομική Ευθύνη (Medical Liability)

Το σύστημα **δεν αποτελεί ιατροτεχνολογικό προϊόν** (Medical Device) υπό την έννοια του MDR (EU 2017/745), καθώς δεν εκδίδει διαγνώσεις αλλά ενδεικτικές εκτιμήσεις, περιλαμβάνει ρητή δήλωση (Disclaimer) ότι «Αυτό το εργαλείο προσφέρει ενδεικτική αξιολόγηση και δεν αντικαθιστά επαγγελματία ψυχικής υγείας», και διαθέτει ευρετικό επίπεδο ελέγχου (Regex) ως επιπλέον μηχανισμό.

Θα πρέπει να σημειωθεί ότι ο χαρακτηρισμός ενός προϊόντος ως ιατροτεχνολογικό ή μη απαιτεί νομική αξιολόγηση από εξειδικευμένους φορείς και δεν μπορεί να καθοριστεί μόνο από τους σχεδιαστές του συστήματος. Η παρούσα περιγραφή αποτελεί την πρόθεση σχεδιασμού και όχι νομική γνωμοδότηση.

7.2.γ Διαφάνεια και Ενήμερη Συγκατάθεση

Το σύστημα σχεδιάστηκε με στόχο τη συμμόρφωση με το Άρθρο 13 GDPR [4]. Οι χρήστες ενημερώνονται ρητά πως τα δεδομένα τους επεξεργάζονται από ΑΙ, παρέχεται επιλογή opt-out για τη συλλογή ανώνυμων δεδομένων βελτίωσης, και η αρχιτεκτονική Privacy by Design επιδιώκει την αποφυγή μόνιμης αποθήκευσης προσωπικών δεδομένων.

7.3 Περιορισμοί της Έρευνας

Η εργασία παρουσιάζει τους ακόλουθους περιορισμούς, οι οποίοι πρέπει να ληφθούν υπόψη κατά την ερμηνεία των αποτελεσμάτων:

7.3.α Συνθετικά vs Πραγματικά Δεδομένα

Περιορισμός: Τα δεδομένα εκπαίδευσης παράχθηκαν τεχνητά μέσω του αλγορίθμου Gaussian Personas. Αν και η μέθοδος βασίζεται σε στατιστικές αρχές [23], δεν μπορεί να αντικαταστήσει την πολυπλοκότητα του φυσικού λόγου (θόρυβος, αταξία, σπάνιες φράσεις).

Επίδραση: Το χαμηλό R^2 (~0.30) ενδέχεται να αντανακλά τόσο την εσκευμένη στοχαστικότητα του αλγορίθμου όσο και περιορισμούς του μοντέλου. Η πραγματική απόδοση σε κλινικά δεδομένα παραμένει άγνωστη.

Μετριασμός: Η πιλοτική χρήση του συστήματος θα επιτρέψει την επανεκπαίδευση σε real-world data (βλ. Ενότητα 7.4).

7.3.β Έλλειψη Κλινικής Επικύρωσης

Περιορισμός: Η αξιολόγηση βασίστηκε σε αυτόματες μετρικές (RMSE, MAE) και όχι σε σύγκριση με αξιολογήσεις από επαγγελματίες ψυχικής υγείας.

Επίδραση: Δεν μπορεί να αποδειχθεί η οικολογική εγκυρότητα (ecological validity) του συστήματος σε πραγματικές συνθήκες.

Μετριασμός: Προτείνεται συνεργασία με Πανεπιστημιακό Συμβουλευτικό Κέντρο για prospective validation study.

7.3.γ Μονογλωσσικό Μοντέλο

Περιορισμός: Το σύστημα υποστηρίζει αποκλειστικά την Ελληνική γλώσσα. Δεν μπορεί να χειριστεί code-switching (π.χ. Ελληνικά με Αγγλικούς όρους) ή πολύγλωσσους χρήστες.

Επίδραση: Περιορίζεται η εφαρμογή σε ομογενείς ελληνόφωνους πληθυσμούς.

Μετριασμός: Μελλοντική χρήση multilingual BERT (XLM-RoBERTa) θα μπορούσε να επεκτείνει τη λειτουργικότητα.

7.3.δ Απουσία Χρονικής Μοντελοποίησης

Περιορισμός: Το σύστημα αναλύει μεμονωμένα κείμενα (snapshot), χωρίς να παρακολουθεί την εξέλιξη των συμπτωμάτων στον χρόνο.

Επίδραση: Δεν μπορεί να εντοπίσει τάσεις ή να προβλέψει αλλαγές.

Μετριασμός: Προσθήκη temporal modeling (π.χ. LSTM για διαχρονική ανάλυση) σε μελλοντικές εκδόσεις.

7.4 Μελλοντικές Επεκτάσεις (Future Work)

7.4.α Μετάβαση σε Δεδομένα Πραγματικής Χρήσης

Ένας βασικός στόχος για τη μελλοντική εξέλιξη του συστήματος είναι η μετάβαση από τα συνθετικά δεδομένα σε δεδομένα πραγματικής χρήσης (Real-world Data). Καθώς τα συνθετικά δεδομένα στερούνται συχνά της αταξίας του φυσικού λόγου, η επόμενη φάση θα μπορούσε να αφορά την πιλοτική διάθεση της εφαρμογής σε ελεγχόμενο περιβάλλον χρηστών.

Η προτεινόμενη διαδικασία θα μπορούσε να ακολουθήσει το μοντέλο Human-in-the-loop (HITL) learning. Αυτό θα περιλαμβάνει πιλοτική διάθεση σε μικρό αριθμό φοιτητών με συναίνεση για ανώνυμη συλλογή δεδομένων, επισημάνση ποιότητας από ψυχολόγους με έλεγχο inter-rater reliability, επαναληπτική εκπαίδευση του BERT με συνδυασμό συνθετικών και πραγματικών δεδομένων, και σύγκριση απόδοσης μεταξύ των διαφορετικών μοντέλων.

Αυτός ο κύκλος ανάδρασης θα μπορούσε να συμβάλει στην επίτευξη οικολογικής εγκυρότητας (Ecological Validity), επιδιώκοντας το μοντέλο να λειτουργεί αξιόπιστα σε πραγματικές συνθήκες, καθώς και στη μείωση του χάσματος προσαρμογής (Domain Adaptation) μεταξύ των στατιστικών κατανομών των συνθετικών δεδομένων και της πραγματικής χρήσης.

7.4.β Multimodal Features & Φωνητική Ανάλυση

Η ενσωμάτωση φωνητικών χαρακτηριστικών (prosody, pitch, tempo) θα μπορούσε να βελτιώσει την ανίχνευση σαρκασμού και συναισθηματικής έντασης. Μελέτες υποδεικνύουν ότι ο συνδυασμός NLP + audio features μπορεί να αυξήσει την ακρίβεια [57].

7.4.γ Edge AI & On-Device Inference

Για την ενίσχυση της προστασίας δεδομένων και την εξάλειψη της ανάγκης για server, προτείνεται η μετατροπή του μοντέλου σε quantized BERT (8-bit) για εκτέλεση εντός της συσκευής Android [37]. Αυτό επιφέρει trade-offs: ως πλεονέκτημα, zero-latency και πλήρης offline λειτουργία, αλλά ως μειονέκτημα, μείωση ακρίβειας κατά ~3-5%.

7.4.δ Ενσωμάτωση Τεχνικών Explainable AI (XAI)

Όπως αναφέρθηκε στην Ενότητα 2.4, τεχνικές όπως SHAP και LIME θα μπορούσαν να παρέχουν εξηγήσεις για τις προβλέψεις του μοντέλου. Η ενσωμάτωση τέτοιων τεχνικών θα αύξανε τη διαφάνεια του συστήματος και θα διευκόλυνε τον έλεγχο από επαγγελματίες.

ΚΕΦΑΛΑΙΟ 8: ΣΥΜΠΕΡΑΣΜΑΤΑ

Η παρούσα εργασία παρουσιάζει τον σχεδιασμό και την ανάπτυξη ενός πρωτοτύπου συστήματος mHealth, το οποίο επιχειρεί να αντιμετωπίσει τις προκλήσεις της έλλειψης δεδομένων και της γλωσσικής υποστήριξης για την Ελληνική γλώσσα στον τομέα της ψυχικής υγείας.

8.1 Κύρια Χαρακτηριστικά

Η εργασία περιλαμβάνει τη Μεθοδολογία Συνθετικών Δεδομένων, όπου η μέθοδος των Gaussian Personas επέτρεψε την εκπαίδευση μοντέλων χωρίς χρήση πραγματικών δεδομένων ασθενών, παράγοντας δεδομένα που προσομοιώνουν μοτίβα απαντήσεων. Θα πρέπει να σημειωθεί ότι η μέθοδος αυτή αποτελεί απλοποίηση και δεν έχει επικυρωθεί ως προς την οικολογική της εγκυρότητα με σύγκριση με πραγματικά κλινικά δεδομένα.

Επίσης, η εργασία προτείνει την Υβριδική Αρχιτεκτονική, όπου ο συνδυασμός BERT + Regex Safety Layer στοχεύει στην ανίχνευση γλωσσικών ενδείξεων (RMSE ~0.74 στα συνθετικά δεδομένα), ενώ ο ευρετικός μηχανισμός Regex προσθέτει ένα επιπλέον επίπεδο ελέγχου για προκαθορισμένες λέξεις-κλειδιά. Θα πρέπει να τονιστεί ότι η αναφερόμενη απόδοση (Recall 99% για το Regex layer) αφορά αποκλειστικά συνθετικά δεδομένα και δεν έχει επικυρωθεί σε πραγματικές συνθήκες.

Ακόμη, η εργασία παρέχει Υποστήριξη Ελληνικής Γλώσσας, αναπτύσσοντας ένα σύστημα αξιολόγησης ανοικτού κώδικα για την Ελληνική γλώσσα, ευθυγραμμισμένο με τυποποιημένες κλίμακες (BDI-II, SCL-90). Αποτελεί μια προσπάθεια κάλυψης του ερευνητικού κενού για τις γλώσσες χαμηλών πόρων.

Τέλος, σε ό,τι αφορά την Πρακτική Υλοποίηση, ο χρόνος απόκρισης ~70ms καθιστά το σύστημα κατάλληλο για εφαρμογές πραγματικού χρόνου, ενώ η αρχιτεκτονική Privacy by Design επιδιώκει την προστασία της ιδιωτικότητας των χρηστών.

8.2 Περιορισμοί και Επισημάνσεις

Η εργασία έχει σημαντικούς περιορισμούς που πρέπει να ληφθούν υπόψη. Δεν έχει πραγματοποιηθεί κλινική επικύρωση και όλες οι μετρικές προέρχονται από συνθετικά δεδομένα. Ο ευρετικός μηχανισμός Regex δεν μπορεί να εντοπίσει έμμεσες ή μεταφορικές εκφράσεις κινδύνου. Η απλοποίηση Gaussian δεν αποδίδει την πλήρη πολυπλοκότητα του φυσικού λόγου. Η συμμόρφωση με GDPR απαιτεί νομική αξιολόγηση που δεν έχει πραγματοποιηθεί.

8.3 Προοπτική Συνεισφορά

Η εργασία θα μπορούσε να συνεισφέρει στη διεθνή έρευνα σε τρία επίπεδα. Σε θεωρητικό επίπεδο, προσαρμόζει τη μεθοδολογία διπλού ελέγχου [30] από το πεδίο της όρασης υπολογιστών στο πεδίο της Επεξεργασίας Φυσικής Γλώσσας. Σε μεθοδολογικό επίπεδο, διερευνά τα συνθετικά δεδομένα ως πιθανή λύση για τη data scarcity σε ρυθμιζόμενα περιβάλλοντα. Σε πρακτικό επίπεδο, παρέχει κώδικα ανοικτού κώδικα που θα μπορούσε να προσαρμοστεί σε άλλες γλώσσες και κλίμακες.

8.4 Από το Proof-of-Concept στην Πιθανή Κλινική Εφαρμογή

Η παρούσα υλοποίηση λειτουργεί ως τεχνική απόδειξη σκοπιμότητας (Proof-of-Concept). Η ουσιαστική αξία της έρευνας θα καθοριστει από τη δυνατότητα προσαρμογής του συστήματος μέσω δεδομένων πεδίου (in-the-wild data). Όπως καταδείχθηκε στην Ενότητα 7.3, τα συνθετικά δεδομένα θέτουν όρια στην προβλεπτική ικανότητα.

Η επόμενη φάση θα ήταν η πιλοτική διάθεση σε πραγματικό περιβάλλον (π.χ. Πανεπιστημιακό Συμβουλευτικό Κέντρο) υπό επίβλεψη επαγγελματιών ψυχικής υγείας, με τις απαραίτητες εγκρίσεις Επιτροπής Ηθικής. Η μετάβαση σε ένα μοντέλο συνεχούς μάθησης (continuous learning) με Human-in-the-Loop validation θα μπορούσε να επιτρέψει τη βελτίωση της ακρίβειας μέσω domain adaptation σε πραγματικά δεδομένα χρηστών, την ανίχνευση νέων μοτίβων που δεν υπάρχουν στα συνθετικά δεδομένα, και την οικοδόμηση εμπιστοσύνης μέσω διαφανούς συνεργασίας με την επιστημονική κοινότητα.

8.5 Τελικό Μήνυμα

Το προτεινόμενο σύστημα δεν φιλοδοξεί να αντικαταστήσει τον επαγγελματία ψυχικής υγείας και δεν είναι κατάλληλο για τέτοια χρήση στην τρέχουσα μορφή του. Εάν επικυρωθεί κλινικά σε μελλοντική εργασία, θα μπορούσε ενδεχομένως να λειτουργήσει ως εργαλείο πρώιμης υποστήριξης, βοηθώντας στον εντοπισμό φοιτητών που ενδέχεται να χρειάζονται βοήθεια και παραπέμποντάς τους στις κατάλληλες δομές. Μέχρι τότε, παραμένει ερευνητικό πρωτότυπο.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] MDPI. (2024). Psychosocial Complaints and Life Satisfaction Among Greek University Students: A Pre- and Pandemic-Era Comparison. *European Journal of Investigation in Health, Psychology and Education*, 4(11). <https://www.mdpi.com/2673-8112/4/11/122>
- [2] MDPI. (2024). Mediterranean Diet Adherence Is Associated with Lower Prevalence of Depression and Anxiety in University Students: A Cross-Sectional Study in Greece. *Medical Sciences*, 14(1). <https://www.mdpi.com/2079-9721/14/1/19>
- [3] I-JMR. (2024). Screening for Depression Using Natural Language Processing: Literature Review. *Interactive Journal of Medical Research*, 13. <https://www.ijmr.org/2024/1/e55067>
- [4] El Emam, K., Mosquera, L., & Bass, J. (2020). Evaluating identity disclosure risk in synthetic health data. *Journal of the American Medical Informatics Association*, 27(3), 471–481. <https://doi.org/10.1093/jamia/ocz221>
- [5] MDPI. (2024). A Systematic Review of Synthetic Data Generation Techniques Using Generative AI. *Electronics*, 13(17). <https://www.mdpi.com/2079-9292/13/17/3509>
- [6] NIH/PMC. (2021). Prevalence and associated factors of anxiety and depression in students at a Greek university during COVID-19 lockdown. *Public Health in Practice*, 2. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8477232/>
- [7] NIH/PMC. (2024). Prevalence of depressive tendencies among college students and the influence of attributional styles on depressive tendencies in the post-pandemic era. *Frontiers in Psychology*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10850216/>
- [8] Vasiou, A., et al. (2022). Stress, Anxiety and Depression Prevalence among Greek University Students during COVID-19 Pandemic: A Two-Year Survey. *International Journal of Environmental Research and Public Health*, 19(15). <https://doi.org/10.3390/ijerph19159679>
- [9] PubMed. (2025). Exploring the Levels of Stress, Anxiety, Depression, Resilience, Hope, and Spiritual Well-Being Among Greek Dentistry and Nursing Students. *PubMed*. <https://pubmed.ncbi.nlm.nih.gov/39791661/>
- [10] NIH/PMC. (2024). Effectiveness of artificial intelligence chatbots on mental health & well-being in college students: a rapid systematic review. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12582922/>
- [11] NIH/PMC. (2025). The state of mental health in Greece: An international comparative analysis using data from the Global Mental Health Countdown 2030. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12171043/>
- [12] BMJ Open. (2025). Students' perceptions of AI mental health chatbots: an exploratory qualitative study at Sultan Qaboos University. *BMJ Open*, 15(10). <https://bmjopen.bmj.com/content/15/10/e103893>

- [13] European Data Protection Board. (2025). *AI Privacy Risks & Mitigations – Large Language Models (LLMs)*. EDPB. <https://www.edpb.europa.eu/system/files/2025-04/ai-privacy-risks-and-mitigations-in-llms.pdf>
- [14] NIH/PMC. (2024). Effectiveness of a Web-based and Mobile Therapy Chatbot on Anxiety and Depressive Symptoms in Subclinical Young Adults: Randomized Controlled Trial. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10993129/>
- [15] Provoost, S., et al. (2020). Conversational agents and the making of mental health recovery. *ResearchGate*. https://www.researchgate.net/publication/347077836_Conversational_agents_and_the_making_of_mental_health_recovery
- [16] JMIR. (2025). Expert and Interdisciplinary Analysis of AI-Driven Chatbots for Mental Health Support: Mixed Methods Study. *Journal of Medical Internet Research*. <https://www.jmir.org/2025/1/e67114>
- [17] NIH/PMC. (2024). Effectiveness of a Mental Health Chatbot for People With Chronic Diseases: Randomized Controlled Trial. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11176869/>
- [18] NIH/PMC. (2025). MentalRoBERTa-Caps: A capsule-enhanced transformer model for mental health classification. *Scientific Reports*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12284574/>
- [19] ACL Anthology. (2024). Establishing Control Corpora for Depression Detection in Modern Greek: Methodological Insights. *LREC-COLING 2024*. <https://aclanthology.org/2024.rapid-1.8.pdf>
- [20] Koutsikakis, A., & Chalkidis, I. (2020). GREEK-BERT: The Greeks visiting Sesame Street. *arXiv preprint arXiv:2008.12014*. <https://arxiv.org/abs/2008.12014>
- [21] MDPI. (2023). A Comparative Study of Sentiment Classification Models for Greek Reviews. *Information*, 8(9). <https://www.mdpi.com/2504-2289/8/9/107>
- [22] NIH/PMC. (2025). Analyzing student mental health with RoBERTa-Large: a sentiment analysis and data analytics approach. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12575187/>
- [23] NIH/PMC. (2025). Large language models generating synthetic clinical datasets: a feasibility and comparative analysis with real-world perioperative data. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11836953/>
- [24] Frontiers. (2025). Synthetic data in medical imaging within the EHDS: a path forward for ethics, regulation, and standards. *Frontiers in Digital Health*. <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2025.1620270/full>
- [25] JMIR AI. (2025). Utility-based Analysis of Statistical Approaches and Deep Learning Models for Synthetic Data Generation With Focus on Correlation Structures. *JMIR AI*. <https://ai.jmir.org/2025/1/e65729/>

- [26] Abd-Alrazaq, A., et al. (2021). Artificial Intelligence–Based Chatbot for Anxiety and Depression in University Students: Pilot Randomized Controlled Trial. *JMIR Formative Research*, 5(8), e20678. <https://doi.org/10.2196/20678>
- [27] Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433–460. <https://courses.cs.umbc.edu/471/papers/turing.pdf>
- [28] Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2021). Activation functions in deep learning: A comprehensive survey and benchmark. *arXiv preprint arXiv:2109.14545*. <https://arxiv.org/abs/2109.14545>
- [29] Paszke, A., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems (NeurIPS)*. <https://arxiv.org/abs/1912.01703>
- [30] Βιολέντης, Α. (2025). *Ανίχνευση Διαταραχών Γραφής και Ορθογραφίας με Χρήση Τεχνητής Νοημοσύνης*. (Πτυχιακή Εργασία). Πανεπιστήμιο Θεσσαλίας. <https://ir.lib.uth.gr/>
- [31] Vaswani, A., et al. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*. <https://arxiv.org/abs/1706.03762>
- [32] zrok. (n.d.). *Getting Started / Documentation*. Retrieved from <https://docs.zrok.io/docs/getting-started/>
- [33] Hugging Face. (n.d.). *nlpaueb/bert-base-greek-uncased-v1 (Model Card)*. Retrieved from <https://huggingface.co/nlpaueb/bert-base-greek-uncased-v1>
- [34] FastAPI. (n.d.). *Concurrency and async/await documentation*. Retrieved from <https://fastapi.tiangolo.com/async/>
- [35] Ramalho, L. (2022). *Fluent Python: Clear, Concise, and Effective Programming* (2nd ed.). O'Reilly Media.
- [36] Python Software Foundation. (n.d.). *multiprocessing — Process-based parallelism*. Retrieved from <https://docs.python.org/3/library/multiprocessing.html>
- [37] Dettmers, T., Pagnoni, A., Holtzman, A., & Zettlemoyer, L. (2023). QLoRA: Efficient finetuning of quantized LLMs. *Advances in Neural Information Processing Systems (NeurIPS)*. <https://doi.org/10.48550/arXiv.2305.14314>
- [38] Android Developers. (n.d.). *ANDROID_ID, security best practices & certificate pinning guidance*. Retrieved from <https://developer.android.com/reference/android/provider/Settings.Secure>
- [39] Uvicorn. (n.d.). *ASGI concepts & FastAPI deployment patterns*. Retrieved from <https://uvicorn.dev/concepts/asgi/>
- [40] Whisman, M. A., & Richardson, E. D. (2015). Normative Data on the Beck Depression Inventory-Second Edition (BDI-II) in College Students. *Journal of Clinical Psychology*, 71(9), 898-907. <https://doi.org/10.1002/jclp.22188>

- [41] Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.
- [42] de Sá Junior, A. R., et al. (2019). Can Gender and Age Impact on Response Pattern of Depressive Symptoms Among College Students? A Differential Item Functioning Analysis. *Frontiers in Psychiatry*. <https://doi.org/10.3389/fpsy.2019.00050>
- [43] Derogatis, L. R. (1994). *SCL-90-R: Administration, Scoring and Procedures Manual*. National Computer Systems.
- [44] Gempp, R., & Avendaño, C. (2008). Normative data and psychometric properties of the SCL-90-R in Chilean College students. *Terapia Psicológica*, 26(1), 39-58. https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-48082008000100003
- [45] Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy via chatbot. *JMIR Mental Health*, 4(2), e19. <https://doi.org/10.2196/mental.7759>
- [46] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. <https://doi.org/10.1145/2939672.2939778>
- [47] MDPI. (2023). Explainable Artificial Intelligence for Workplace Mental Health. *Systems*, 12(4). <https://www.mdpi.com/2227-9709/12/4/130>
- [48] NIH/PMC. (2025). Explainable AI for Depression Detection and Severity Classification. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12425426/>
- [49] MDPI. (2024). The Black Box Paradox: AI Models and the Epistemological Crisis in Motor Control Research. *Information*, 16(10). <https://www.mdpi.com/2078-2489/16/10/823>
- [50] NIH/PMC. (2024). Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10997568/>
- [51] NIH/PMC. (2025). Explainable AI for Clinical Outcome Prediction: A Survey of Clinician Perspectives. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12150750/>
- [52] Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational AI agent. *The Lancet Digital Health*, 1(1), e52-e53. [https://doi.org/10.1016/S2589-7500\(18\)30025-0](https://doi.org/10.1016/S2589-7500(18)30025-0)
- [53] NIH/PMC. (2025). Ethical Decision-Making Guidelines for Mental Health Clinicians in the Artificial Intelligence (AI) Era. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12692113/>
- [54] London, A. J. (2017). *Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability*. Hastings Center Report, Carnegie Mellon University. <https://www.cmu.edu/dietrich/philosophy/docs/london/hastings.pdf>

- [55] Clark, K., et al. (2019). What Does BERT Look At? An Analysis of BERT's Attention. *Proceedings of the 2019 ACL Workshop BlackboxNLP*. <https://aclanthology.org/W19-4828.pdf>
- [56] ResearchGate. (2022). Exploring Transformer-Based Learning for Negation Detection in Biomedical Texts. *ResearchGate*. https://www.researchgate.net/publication/362625708_Exploring_Transformer-Based_Learning_for_Negation_Detection_in_Biomedical_Texts
- [57] ResearchGate/MDPI. (2024). RAMHA: A Hybrid Social Text-Based Transformer with Adapter for Mental Health Emotion Classification. *Electronics*. <https://www.mdpi.com/2079-9292/12/4/987>
- [58] arXiv. (2025). AttentionDep: Domain-Aware Attention for Explainable Depression Severity Assessment. *arXiv*. <https://arxiv.org/html/2510.00706v1>
- [59] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*. <https://aclanthology.org/N19-1423/>
- [60] Touvron, H., Martin, L., Stone, K., et al. (2023). Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*. <https://arxiv.org/abs/2307.09288>
- [61] Kovaleva, O., et al. (2019). Revealing the Dark Secrets of BERT. *arXiv preprint arXiv:1908.08593*. <https://arxiv.org/abs/1908.08593>
- [62] Rogers, A., et al. (2020). A Primer in BERTology: What We Know About How BERT Works. *Transactions of the Association for Computational Linguistics*. <https://arxiv.org/abs/2002.12327>
- [63] Michel, P., et al. (2019). Are Sixteen Heads Really Better than One? *Advances in Neural Information Processing Systems (NeurIPS)*. <https://arxiv.org/abs/1905.10650>
- [64] arXiv. (2025). User Perception of Attention Visualizations: Effects on Interpretability Across Evidence-Based Medical Documents. *arXiv*. <https://arxiv.org/html/2508.10004v1>
- [65] MDPI. (2024). The Promise of Explainable AI in Digital Health for Precision Medicine: A Systematic Review. *Journal of Personalized Medicine*. <https://www.mdpi.com/2075-4426/14/3/277>
- [66] NIH/PMC. (2021). Artificial Intelligence for Mental Healthcare: Clinical Applications, Barriers, Facilitators, and Artificial Wisdom. *Frontiers in Psychiatry*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8349367/>
- [67] MDPI. (2024). Ethical Considerations in Artificial Intelligence Interventions for Mental Health and Well-Being. *Social Sciences*. <https://www.mdpi.com/2076-0760/13/7/381>

- [68] NIH/PMC. (2020). Illuminating the Black Box: Interpreting Deep Neural Network Models for Psychiatric Research. *Frontiers in Psychiatry*.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC7658441/>
- [69] NIH/PMC. (2022). Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence*.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC9122117/>
- [70] NIH/PMC. (2025). Ethical decision-making for AI in mental health: the Integrated Ethical Approach. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12315656/>
- [71] NIH/PMC. (2025). Artificial intelligence in mental health care: a systematic review. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12017374/>
- [72] MDPI. (2023). An Attention-Based BERT–CNN–BiLSTM Model for Depression. *Information*. <https://www.mdpi.com/2504-2289/9/12/310>
- [73] Abd-Alrazaq, A. A., et al. (2020). Effectiveness and safety of conversational agents in mental health. *Journal of Medical Internet Research*, 22(7), e16021.
<https://doi.org/10.2196/16021>
- [74] Frontiers. (2024). Differing perspectives on artificial intelligence in mental healthcare among patients. *Frontiers in Digital Health*, 6.
<https://doi.org/10.3389/fdgth.2024.1410758>
- [75] NIH/PMC. (2024). Patient Perspectives on AI for Mental Health Care: Cross-Sectional Survey Study. *PMC*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11447436/>
- [76] Gorelick, M., & Ozsvald, I. (2020). *High Performance Python: Practical Performant Programming for Humans* (2nd ed.). O'Reilly Media.
- [77] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*.
<https://arxiv.org/abs/1412.6980>
- [78] Loshchilov, I., & Hutter, F. (2019). Decoupled weight decay regularization. *International Conference on Learning Representations (ICLR)*.
<https://arxiv.org/abs/1711.05101>
- [79] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
<http://www.deeplearningbook.org>
- [80] NeurIPS. (2024). Why Do We Need Weight Decay in Modern Deep Learning? *Proceedings of NeurIPS*. <https://arxiv.org/abs/2310.04415>
- [81] NeurIPS. (2024). Why Do Transformers Need Adam: A Hessian Perspective. *Proceedings of NeurIPS*. <https://arxiv.org/abs/2402.16788>
- [82] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>

- [83] Scholastica. (n.d.). Determining a Meaningful R-squared Value in Clinical Medicine. *Academic Med Surg*. <https://academic-med-surg.scholasticahq.com/article/125154>
- [84] NIH/PMC. (2015). Interpreting regression models in clinical outcome studies. *BMJ*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4678365/>
- [85] NIH/PMC. (2015). Minimal clinically important difference on the Beck Depression Inventory-II according to the patient's perspective. *Journal of Affective Disorders*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4611356/>
- [86] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: with Applications in R* (2nd ed.). Springer. <https://www.statlearning.com/>
- [87] NIH/PMC. (2021). A direct comparison of theory-driven and machine learning prediction of suicide: A meta-analysis. *PLOS ONE*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8041204/>
- [88] Μπαλκάμος, Σ. (2026). *MentalHealth-Chatbot-Greek: Σύστημα Αξιολόγησης Ψυχικής Υγείας με χρήση Τεχνητής Νοημοσύνης*. GitHub Repository. Διαθέσιμο στο: <https://github.com/sbalkamos/MentalHealth-Chatbot-Greek>