

# Η συνεισφορά της Τεχνητής Νοημοσύνης στην πρώιμη διάγνωση της νόσου Alzheimer μέσω χαρακτηριστικών του λόγου

Φωτεινή Σαββίδου, AEM: 9657, sfoteini@ece.auth.gr

## Περίληψη

Η νόσος Alzheimer είναι ένα από τα κυριότερα νευροεκφυλιστικά νοσήματα και χαρακτηρίζεται από εξασθένιση των νοητικών και γνωστικών λειτουργιών, συμπεριφορικές και ψυχολογικές διαταραχές και δυσκολία εκτέλεσης καθημερινών δραστηριοτήτων. Ένα σύνηθες σύμπτωμα της νόσου είναι η διαταραχή της ομιλίας, που εκδηλώνεται με τη μείωση της ευφράδειας, την απώλεια σημασιολογικής ευχέρειας, την αδυναμία κατανόησης του λόγου και την αλλοίωση της προσωδίας καθιστώντας δύσκολη την επικοινωνία. Ο λόγος είναι μια πολύτιμη πηγή κλινικών πληροφοριών στη νόσο Alzheimer, καθώς η γλωσσική έκπτωση εμφανίζεται από τα πρώτα στάδια της νόσου. Συνεπώς, τα χαρακτηριστικά της ομιλίας, ως μέσο για έγκαιρη διάγνωση αλλά και αξιολόγηση της εξέλιξης της νόσου, έχουν μελετηθεί εκτενώς. Στόχος της παρούσας εργασίας είναι η μελέτη προηγούμενων ερευνών σχετικά με τη χρήση της τεχνητής νοημοσύνης για την πρώιμη διάγνωση της νόσου Alzheimer από χαρακτηριστικά του λόγου, ο προσδιορισμός των γλωσσικών, παραγλωσσικών και προσωδιακών χαρακτηριστικών που μπορούν να χρησιμοποιηθούν ως βιοδείκτες, η αξιολόγηση της αποτελεσματικότητας των τεχνικών μηχανικής μάθησης για τη διάγνωση της νόσου σε πρώιμο στάδιο και η επισήμανση των πιθανών προβλημάτων και αδυναμιών που εντοπίστηκαν.

**Λέξεις-κλειδιά:** Νόσος Alzheimer, Διαταραχές του λόγου, Ανάλυση ομιλίας, Τεχνητή νοημοσύνη

## 1 Εισαγωγή

Η νόσος Alzheimer (AD) είναι μία χρόνια νευροεκφυλιστική ασθένεια του Κεντρικού Νευρικού Συστήματος που χαρακτηρίζεται από εξασθένιση των νοητικών και γνωστικών λειτουργιών, συμπεριφορικές και ψυχολογικές διαταραχές και δυσκολία εκτέλεσης καθημερινών δραστηριοτήτων. Αποτελεί ένα από τα πιο συνήθη αίτια της άνοιας. Τα συμπτώματα της νόσου εμφανίζονται σταδιακά και επιδεινώνονται με την πάροδο του χρόνου.<sup>[1]</sup> Σήμερα, υπολογίζεται πως η νόσος Alzheimer επηρεάζει περίπου το 6% του πληθυσμού ηλικίας άνω των 65 ετών<sup>[1]</sup> και συνολικά πάνω από 50 εκατομμύρια ανθρώπους παγκοσμίως<sup>[2]</sup>. Αν και η ταχύτητα της εξέλιξης ποικίλλει, το μέσο προσδόκιμο επιβίωσης μετά τη διάγνωση είναι τέσσερα έως οκτώ έτη.<sup>[4]</sup> Η αιτία της νόσου Alzheimer δεν είναι γνωστή. Ωστόσο, επιστημονικές έρευνες αποδεικνύουν πως πολλοί παράγοντες

συμβάλλουν στην εμφάνιση της πάθησης, όπως ηλικία, κληρονομικότητα, τραυματισμοί στο κεφάλι, κατάθλιψη, σύνδρομο Down και ασθένειες που προκαλούν χρωμοσωμικές μεταλλάξεις. Ακόμη, ο διαβήτης, η υπέρταση, η υψηλή χοληστερόλη, η μειωμένη σωματική και γνωστική δραστηριότητα, αλλά και ο γενικότερος τρόπος ζωής περιλαμβάνονται στους παράγοντες που αυξάνουν τον κίνδυνο εμφάνισης της νόσου.<sup>[1],[5]</sup>

Στα συνήθη πρώιμα συμπτώματα της AD συγκαταλέγεται η εξασθένηση των γνωστικών λειτουργιών, όπως η απώλεια μνήμης (αδυναμία των ασθενών να θυμηθούν καταστάσεις, πρόσωπα και ονόματα), οι διαταραχές του λόγου και η σταδιακή μείωση της γλωσσικής ευφράδειας, η απάθεια και η κατάθλιψη. Καθώς η ασθένεια εξελίσσεται, οι ασθενείς παρουσιάζουν μειωμένη λεκτική ευχέρεια και γλωσσικές δεξιότητες, αδυναμία επικοινωνίας, προβλήματα στον προσανατολισμό και αδυνατούν να εκτελέσουν σύνθετες καθημερινές δραστηριότητες, ενώ οι συμπεριφορικές μεταπτώσεις και οι ψυχικές διαταραχές εντείνονται. Οι ασθενείς εμφανίζουν ψευδαισθήσεις, αυταπάτες, σύγχυση και άγχος.<sup>[1],[4]</sup>

Τα παραπάνω συμπτώματα είναι απόρροια της απώλειας των νευρώνων, της διαταραχής της επικοινωνίας μεταξύ των νευρώνων και της καταστροφής των συνάψεων. Σε πρώτο στάδιο, η νόσος Alzheimer καταστρέφει τους νευρώνες στον ενδορινικό φλοιό και στον ιππόκαμπο, περιοχές του εγκεφάλου που σχετίζονται με τη λειτουργία της μνήμης. Με την εξέλιξη της νόσου, η απώλεια νευρικών κυττάρων επεκτείνεται και στον εγκεφαλικό φλοιό, επηρεάζοντας τις περιοχές του εγκεφάλου που είναι υπεύθυνες για την παραγωγή λόγου, τη λογική και την κοινωνικότητα και οδηγεί στην ατροφία του εγκεφάλου. Έτσι, οι νοητικές, γνωστικές και σωματικές λειτουργίες των ασθενών με νόσο Alzheimer σταδιακά εξασθενούν και αυτοί αδυνατούν να ζήσουν αυτόνομα. Οι αλλοιώσεις που παρατηρούνται στον εγκέφαλο οφείλονται κυρίως στις μη φυσιολογικές συστάδες θραυσμάτων πρωτεΐνης β-αμυλοειδούς (beta-amyloid) που συσσωρεύονται μεταξύ των νευρικών κυττάρων (πλάκες β-αμυλοειδούς) και στην υπέρμετρη συγκέντρωση της πρωτεΐνης Ταυ (Tau) εντός των νευρικών κυττάρων (νευροϊνιδιακοί σωροί).<sup>[4],[6]</sup>

Οι παθολογικές διεργασίες που συντελούνται στον εγκέφαλο και συνδέονται με την AD ξεκινούν αρκετά χρόνια πριν την εμφάνιση των συμπτωμάτων (Προκλινικό στάδιο της AD).<sup>[3]</sup> Στο στάδιο της Ήπιας Γνωστικής Διαταραχής (Mild Cognitive Impairment, MCI), που θεωρείται το πρώιμο στάδιο της AD, οι ασθενείς εμφανίζουν ήπια υποκείμενα συμπτώματα και γνωστικά ελλείμματα, αλλά δεν επηρεάζονται οι δραστηριότητες της καθημερινής ζωής.<sup>[1],[4]</sup> Η έγκαιρη αξιολόγηση της γνωστικής διαταραχής και διάγνωση της AD είναι ιδιαίτερα σημαντική για την επιβράδυνση της εμφάνισης της νόσου και της εκδήλωσης νέων, εντονότερων συμπτωμάτων, αλλά και για την αποτελεσματικότερη παρακολούθηση της πάθησης και παροχή ψυχοκοινωνικής υποστήριξης στους ασθενείς.<sup>[4]</sup> Ο προσδιορισμός λοιπόν των ατόμων με MCI που είναι πιθανό να αναπτύξουν άνοια λόγω της AD είναι ένας κύριος στόχος των πρόσφατων ερευνών.

Στόχος της παρούσας εργασίας είναι η αξιολόγηση της αποτελεσματικότητας των υπολογιστικών μεθόδων επεξεργασίας ομιλίας και των τεχνικών μηχανικής μάθησης για τη διάγνωση της νόσου Alzheimer σε πρώιμο στάδιο και ο προσδιορισμός των χαρακτηριστικών της ομιλίας που παρουσιάζουν ταχεία αλλοίωση λόγω της εξέλιξης της νόσου και μπορούν να χρησιμοποιηθούν ως βιοδείκτες για την έγκαιρη διάγνωσή της. Επιπλέον, η παρούσα μελέτη συνοψίζει τις κύριες μεθόδους που χρησιμοποιούνται για την καταγραφή και τον εντοπισμό των αλλαγών στα χαρακτηριστικά της ομιλίας.

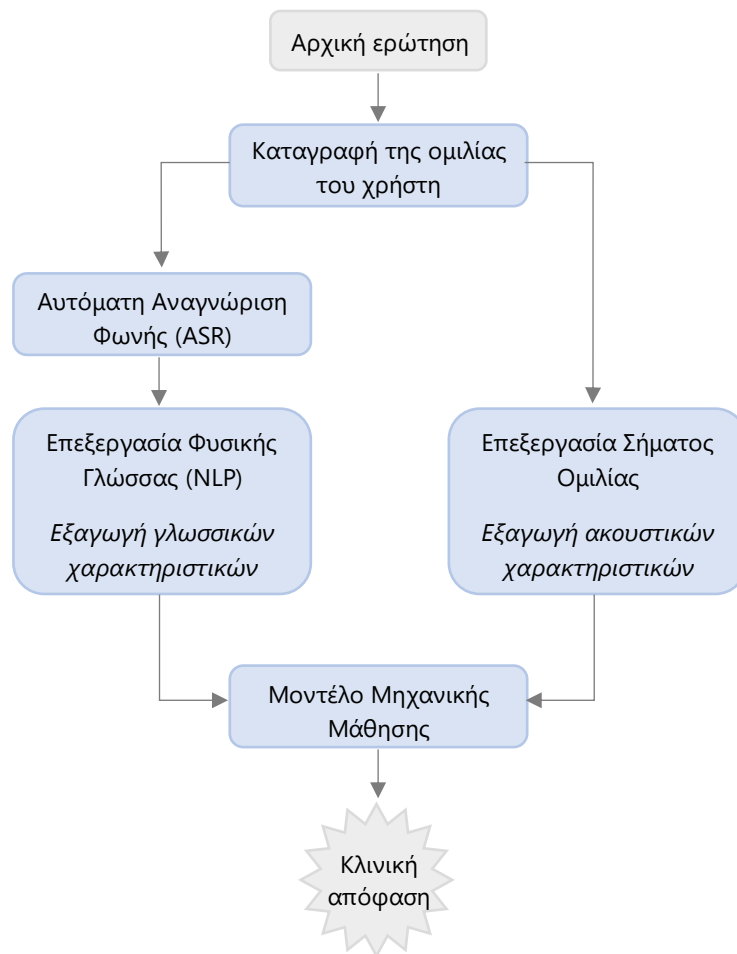
## 2 Επιπτώσεις της νόσου Alzheimer στον λόγο

Αν και η απώλεια μνήμης θεωρείται το κύριο σύμπτωμα της νόσου Alzheimer, η διαταραχή του λόγου εμφανίζεται από τα πρώτα στάδια της AD.<sup>[7]</sup> Η γλωσσική έκπτωση έγκειται στην εξασθένηση του σημασιολογικού και πραγματολογικού επιπέδου της γλώσσας των ασθενών με AD. Σε σημασιολογικό επίπεδο, χαρακτηριστικές επιπτώσεις της AD στον λόγο είναι η δυσκολία εύρεσης και ανάκλησης λέξεων και κατονομασίας ουσιαστικών και ρημάτων, η σημασιολογική παραφασία (επιλογή λανθασμένων λέξεων), η επινόηση λέξεων, η παραγωγή κενού λόγου χωρίς συνοχή και η απώλεια λεκτικής ευχέρειας. Το πραγματολογικό επίπεδο της γλώσσας σχετίζεται με τη χρήση της γλώσσας στο κοινωνικό περιβάλλον και την προσαρμογή της στις κοινωνικές καταστάσεις. Η υπερβολική ομιλία σε ακατάλληλες στιγμές, η πολύ δυνατή ομιλία, η επανάληψη ιδεών και η απόκλιση από το θέμα είναι μερικά συνήθη παραδείγματα διαταραχών του λόγου που εμφανίζονται σε ασθενείς με AD.<sup>[8],[9]</sup>

Η επιδείνωση των νοητικών λειτουργιών και η απουσία λεκτικής ευχέρειας επιφέρει χαρακτηριστικές ακουστικές αλλαγές στην ομιλία των ασθενών με AD και MCI. Η μείωση της ευφράδειας στη διάρκεια της αυθόρμητης ομιλίας μπορεί να ανιχνευτεί στον δισταγμό ενός ασθενή, στην άρθρωση, στην προσωδία και σε αλλαγές στον ρυθμό της ομιλίας. Στους ασθενείς με AD παρατηρείται χαμηλότερος ή κυμαινόμενος ρυθμός ομιλίας, συχνές παύσεις λόγω της δυσκολίας εύρεσης και ανάκλησης λέξεων, βραδύτητα και διακυμάνσεις στη συχνότητα των λέξεων. Σε αρκετές περιπτώσεις, αναφέρονται και μεταπτώσεις στη συναισθηματική κατάσταση του ασθενούς.<sup>[9],[10]</sup>

Πρόσφατες μελέτες σχετικά με τη χρήση της τεχνητής νοημοσύνης (AI) στην έρευνα της AD στοχεύουν στην αξιοποίηση γλωσσικών, παραγλωσσικών και προσωδιακών χαρακτηριστικών της ομιλίας των ασθενών που συλλέγονται με διαφορετικούς τρόπους (για παράδειγμα, καταγραφή του λόγου ασθενών όταν εκτελούν γνωστικές δραστηριότητες ή της αυθόρμητης ομιλίας) και στην εφαρμογή υπολογιστικών μεθόδων επεξεργασίας της ομιλίας για τη έγκαιρη διάγνωση, πρόγνωση ή και παρακολούθηση της εξέλιξης της νόσου.<sup>[7],[9]</sup> Η χρήση χαρακτηριστικών του λόγου και της ομιλίας ως βιοδείκτες για τη νόσο Alzheimer παρουσιάζει ποικίλα πλεονεκτήματα:

- i. Είναι εύκολη η καταγραφή και η συνεχής παρακολούθηση της ομιλίας των ασθενών με τη πάροδο του χρόνου.
- ii. Αποτελεί μη-επεμβατική μέθοδο.
- iii. Η μελέτη γλωσσικών και ακουστικών παραμέτρων της ομιλίας αποτελεί ένα γρήγορο μέσο αξιολόγησης με σχετικά χαμηλή επιβάρυνση για τον ασθενή.
- iv. Οι διαταραχές της ομιλίας και του λόγου εμφανίζονται σε διάφορα στάδια της νόσου, γεγονός που καθιστά τη μελέτη τους έγκυρο μέσο τόσο για την έγκαιρη διάγνωση όσο και για την αξιολόγηση της εξέλιξης της νόσου.
- v. Η παραγωγή του λόγου συνδέεται άρρηκτα με άλλες νοητικές και εκτελεστικές λειτουργίες, όπως η μνήμη, η δυνατότητα ανάκλησης λέξεων, η σημασιολογική ευχέρεια, η άρθρωση και η προσωδία. Έτσι, η μελέτη της ομιλίας αποτελεί ένα ολιστικό μέτρο αξιολόγησης της κατάστασης του ασθενούς.
- vi. Η πρόοδος των μεθόδων επεξεργασίας και ανάλυσης ομιλίας βαίνει αυξανόμενη τις τελευταίες δεκαετίες και ενισχύεται από την εξέλιξη των τεχνικών τεχνητής νοημοσύνης.<sup>[3],[7]</sup>



**Εικόνα 1:** Γενική αρχιτεκτονική υπολογιστικού συστήματος για την εξαγωγή γλωσσικών και ακουστικών χαρακτηριστικών και τη λήψη κλινικών αποφάσεων. Αναδιατύπωση με βάση τους Voleti et al.<sup>[12]</sup>

### 3 Μέθοδοι συλλογής δεδομένων

Τις τελευταίες δεκαετίες, αρκετοί ερευνητές παγκοσμίως επικεντρώνονται στη μελέτη της αυθόρμητης και συνεχούς ομιλίας με στόχο την ανεύρεση γλωσσικών και προσωδιακών χαρακτηριστικών που είναι δυνατό να αξιοποιηθούν για τη διάκριση ασθενών με AD ή MCI σε σχέση με υγιείς ομιλητές. Η καταγραφή των χαρακτηριστικών του λόγου υλοποιείται μέσω κλινικών εξετάσεων και γνωστικών ασκήσεων. Για παράδειγμα, ασκήσεις που έχουν χρησιμοποιηθεί ευρέως σε πρόσφατες μελέτες για την καταγραφή δεδομένων ομιλίας είναι η ανάκληση ιστορίας, η περιγραφή εικόνας, η ανάγνωση αποσπασμάτων κειμένου, η ανάκληση της υπόθεσης ταινίας μικρού μήκους, η επανάληψη μιας ιστορίας, η ανάκληση των συμβάντων της προηγούμενης ημέρας και η αντίστροφη μέτρηση. Εκτός από τις ασκήσεις μονολόγου, πραγματοποιούνται ακόμη (σε μικρότερο ποσοστό) ασκήσεις διαλόγου μέσω της συμμετοχής σε ημι-τυποποιημένη συνέντευξη (δηλαδή, δομημένο διάλογο μεταξύ ασθενούς και επαγγελματία με καθορισμένες ερωτήσεις) και σε συνομιλητική συζήτηση (αυθόρμητος διάλογος μεταξύ δύο συμμετεχόντων).<sup>[7]</sup> Στις έρευνες σχετικά με

την πρώιμη διάγνωση της νοητικής διαταραχής που συνδέεται με την AD αξιοποιούνται διάφορες κατηγορίες γνωστικών ασκήσεων στοχεύοντας στη συλλογή δεδομένων για την αξιολόγηση διαφορετικών εκφάνσεων της γλωσσικής διαταραχής. Μέσω της περιγραφής εικόνας, για παράδειγμα, αξιολογείται η σημασιολογική και η λεκτική ευχέρεια και οι γλωσσικές δεξιότητες (σε επίπεδο μορφολογίας και συντακτικού) του ασθενούς<sup>[3]</sup>, ενώ η δοκιμασία της επανάληψης ιστορίας συντελεί στην εκτίμηση της εξασθένησης της επεισοδιακής και της σημασιολογικής μνήμης.<sup>[11]</sup>



(α)



(β)

**Εικόνα 2:** Ευρέως διαδεδομένες εικόνες που χρησιμοποιούνται ως γνωστικές δοκιμασίες. Η οδηγία προς τους συμμετέχοντες είναι να περιγράψουν όσα βλέπουν στην εικόνα όσο το δυνατόν πιο ολοκληρωμένα. (α) Η δοκιμασία της κλοπής μπισκότου (*The Cookie Theft Picture*) από τη Διαγνωστική Εξέταση της Βοστώνης για την Αφασία. (β) Η σκηνή του πικνίκ (*The Picnic Scene*) από το *Western Aphasia Battery*.<sup>[3],[11]</sup>

Τα χαρακτηριστικά του λόγου που εξάγονται από τις παραπάνω γνωστικές δοκιμασίες διακρίνονται σε δύο κατηγορίες: (α) γλωσσικά χαρακτηριστικά και (β) ακουστικά/προσωδιακά χαρακτηριστικά. Τα γλωσσικά χαρακτηριστικά εξάγονται με μεθόδους Επεξεργασίας Φυσικής Γλώσσας (*Natural Language Processing, NLP*) και περιλαμβάνουν λεξιλογικές και συντακτικές παραμέτρους, όπως η αναλογία τύπου και δείγματος (*type-to-token ratio, TTR*), η πυκνότητα περιεχομένου ή οι δείκτες Yngve και Frazier και σημασιολογικούς και πραγματολογικούς δείκτες. Η παράμετρος TTR είναι ένα μέτρο λεξιλογικής πολυπλοκότητας που υπολογίζεται ως ο λόγος του αριθμού των μοναδικών λέξεων (τύποι, *V*) προς τον συνολικό αριθμό των λέξεων (δείγματα, *N*) σε ένα δεδομένο γλωσσικό παράδειγμα.<sup>[7],[12]</sup>

$$TTR = \frac{V}{N}$$

Ωστόσο, ο δείκτης TTR παρουσιάζει μεροληψία ως προς το μέγεθος του κειμένου, καθώς με την αύξηση του συνολικού αριθμού των λέξεων, ο αριθμός των μοναδικών λέξεων σχεδόν σταθεροποιείται. Για αυτό τον λόγο χρησιμοποιείται ο κινούμενος μέσος όρος τύπου και δείγματος (*moving average type-to-token ratio, MATTR*), ένας δείκτης που στοχεύει στη μείωση της εξάρτησης από το μήκος του κειμένου, υπολογίζοντας το TTR σε ολισθαίνοντα τμήματα του κειμένου.<sup>[12]</sup> Η πυκνότητα περιεχομένου εκφράζει τον αριθμό των πληροφοριών που μεταδίδονται σε μία γλωσσική περίπτωση και υπολογίζεται βάσει της παρακάτω σχέσης. Η υψηλή πυκνότητα περιεχομένου φαί-

νεται πως σχετίζεται με μικρότερη πιθανότητα εμφάνισης της AD, ενώ η χαμηλή πυκνότητα περιεχομένου υποδεικνύει ατροφία του εγκεφάλου.<sup>[7],[12]</sup>

$$CD = \frac{\# \text{ of verbs} + \text{nouns} + \text{adjectives} + \text{adverbs}}{N}$$

Οι ακουστικές παράμετροι που έχουν χρησιμοποιηθεί ευρέως σε πρόσφατες έρευνες περιλαμβάνουν κυρίως τα προσωδιακά χαρακτηριστικά, όπως ο ρυθμός ομιλίας, άρθρωσης και παύσεων και η ένταση της φωνής. Χρησιμοποιούνται ακόμα τα φασματικά χαρακτηριστικά του λόγου, όπως Mel Frequency Cepstral Coefficients (MFCCs), χαρακτηριστικά που σχετίζονται με την ποιότητα της φωνής και ακουστικές παράμετροι που σχετίζονται με την αναγνώριση φωνής, όπως παύσεις, επαναλήψεις, δισταγμοί ή χαρακτηριστικά διαλόγου.<sup>[7],[12]</sup>

Οι König et al. (2015) χρησιμοποίησαν τέσσερις γνωστικές ασκήσεις (αντίστροφη καταμέτρηση από το 305 έως το 285, περιγραφή εικόνας, επανάληψη προτάσεων και σημασιολογική ευχέρεια μέσω κατονομασίας ζώων) για τη συλλογή χαρακτηριστικών του λόγου από υγιείς συμμετέχοντες, ασθενείς με MCI και AD. Τα δεδομένα από τις δύο πρώτες γνωστικές δοκιμασίες αναλύθηκαν ως προς τη συνέχεια του λόγου και υπολογίστηκε η διάρκεια της συνεχόμενης ομιλίας και των παύσεων, η διάρκεια των συνεχόμενων περιοδικών και απεριοδικών τμημάτων του λόγου και στατιστικές παράμετροι αυτών των χαρακτηριστικών με το λογισμικό Praat. Παράμετροι που υπολογίστηκαν από τις άλλες γνωστικές ασκήσεις είναι ο χρόνος φωνητικής αντίδρασης, η σχετική διάρκεια της ομιλίας, το ποσοστό των παύσεων και ο αριθμός των παρατυπιών.<sup>[10]</sup>

Οι Tóth et al. (2018) επικεντρώθηκαν στην ανίχνευση της Ήπιας Γνωστικής Διαταραχής μέσω χαρακτηριστικών της αυθόρμητης ομιλίας κατά την εκτέλεση μνημονικής εργασίας. Για τη συλλογή των δεδομένων χρησιμοποιήθηκε η πειραματική άσκηση της ανάκλησης της υπόθεσης ταινίας μικρού μήκους. Μετά την προβολή μίας ταινίας διάρκειας ενός λεπτού, οι συμμετέχοντες κλήθηκαν να περιγράψουν την υπόθεση της ταινίας (Άσκηση 1) και να αφηγηθούν την προηγούμενη ημέρα τους (Άσκηση 2). Η τρίτη άσκηση αποτελούταν από την προβολή μίας ταινίας και την περιγραφή της υπόθεσής της, αλλά με ένα διάλειμμα διάρκειας ενός λεπτού. Αυτές οι ασκήσεις στοχεύουν στην αξιολόγηση της άμεσης ανάκλησης γεγονότων, της ποιότητας του αυθόρμητου λόγου και της καθυστερημένης ανάκλησης γεγονότων, αντίστοιχα. Για την εξαγωγή των ακουστικών χαρακτηριστικών (διάρκεια πρότασης, ρυθμός ομιλίας και άρθρωσης, αριθμός, συνολική διάρκεια, μέση διάρκεια και ρυθμός παύσεων και σχετική διάρκεια παύσεων ως προς τη συνολική διάρκεια της πρότασης) χρησιμοποιήθηκε το λογισμικό Praat σε συνδυασμό με χειροκίνητη ανάλυση του σήματος, ενώ αναπτύχθηκε ένα ειδικό εργαλείο Αυτόματης Αναγνώρισης Φωνής (Automatic Speech Recognition, ASR) με στόχο την αυτόματη εξαγωγή ακουστικών παραμέτρων-βιοδεικτών.<sup>[9]</sup>

## 4 Τεχνικές Μηχανικής και Βαθιάς Μάθησης

Αλγόριθμοι Μηχανικής Μάθησης (Machine Learning, ML) και Βαθιάς Μάθησης (Deep Learning, DL) χρησιμοποιούνται ευρέως στο τομέα της έρευνας για τη νόσο Alzheimer με στόχο την ανάπτυξη αυτόματων και αξιόπιστων υπολογιστικών συστημάτων για την έγκαιρη διάγνωση της νόσου. Η

αρχιτεκτονική ενός συστήματος μηχανικής μάθησης, που χρησιμοποιείται στην πρόβλεψη της AD με βάση χαρακτηριστικά της ομιλίας, περιγράφεται από τέσσερις παραμέτρους: (1) εργασία μάθησης (learning task), (2) αναπαράσταση δεδομένων (data representation), (3) αλγόριθμος μάθησης (learning algorithm) και (4) μέτρα απόδοσης (performance measures). Όσον αφορά την εργασία μάθησης, μελέτες σχετικά με τη χρήση της τεχνητής νοημοσύνης στην πρώιμη διάγνωση της AD από χαρακτηριστικά του λόγου επικεντρώνονται σε δύο λειτουργίες: την Ταξινόμηση (Classification) και την Παλινδρόμηση (Regression). Οι δύο αυτές εργασίες μάθησης εντάσσονται στον τομέα της επιβλεπόμενης μάθησης (supervised learning), ενώ η τεχνική της Ταξινόμησης συναντάται συχνότερα στις πρόσφατες μελέτες. Με την τεχνική της Ταξινόμησης είναι δυνατή η αντιστοίχιση του δείγματος ομιλίας ενός ασθενή σε μία από τις υπό-μελέτη κατηγορίες, όπως «AD», «MCI» ή «Υγιής», ενώ με τη μέθοδο της Παλινδρόμησης υπολογίζεται μία αριθμητική τιμή η οποία αντιστοιχεί στο αποτέλεσμα μίας ορισμένης νευροψυχολογικής εξέτασης, όπως η Δοκιμασία Σύντομης Γνωστικής Εκτίμησης (Mini-Mental State Examination).<sup>[2],[7]</sup> Η δεύτερη παράμετρος που περιγράφει το σύστημα μηχανικής μάθησης είναι η αναπαράσταση των δεδομένων, η οποία ορίζει τα χαρακτηριστικά της ομιλίας που θα χρησιμοποιηθούν (γλωσσικά ή ακουστικά χαρακτηριστικά) και τον τρόπο κωδικοποίησης των χαρακτηριστικών αυτών. Σε επίπεδο αλγορίθμων, στην πλειονότητα των ερευνών στον τομέα της AD εφαρμόζονται συμβατικοί κατηγοριοποιητές (classifiers), όπως Naive Bayes, Linear SVM, K-Nearest Neighbors και Random Forests και στη συνέχεια αξιολογείται η απόδοσή τους. Στα συνήθη μέτρα απόδοσης περιλαμβάνεται η ακρίβεια (accuracy), η ευαισθησία (sensitivity), η ειδικότητα (specificity) και η χαρακτηριστική καμπύλη λειτουργίας δέκτη (ROC). Λόγω του μικρού μεγέθους των δεδομένων, η χρήση τεχνητών νευρωνικών δικτύων (artificial neural networks) είναι σπανιότερη.<sup>[7],[9]</sup>

Οι König et al. (2015) αξιολόγησαν τρία σενάρια δυαδικής ταξινόμησης: «Υγιής Ομάδα» έναντι «Ομάδα με AD», «Υγιής ομάδα» έναντι «Ομάδα με MCI» και «Ομάδα με AD» έναντι «Ομάδα με MCI». Για κάθε σενάριο ταξινόμησης, χρησιμοποιήθηκε ο έλεγχος Mann-Whitney U test για την επιλογή του βέλτιστου υποσυνόλου των φωνητικών δεδομένων και η ακρίβεια της ταξινόμησης αξιολογήθηκε χρησιμοποιώντας τον ταξινομητή Μηχανής Διανυσμάτων Υποστήριξης (SVM). Στην παρούσα μελέτη επιτεύχθηκε υψηλή ακρίβεια ταξινόμησης ( $\leq 81\%$ ), αποδεικνύοντας πως οι τεχνικές επεξεργασίας ομιλίας παρέχουν αξιόπιστα και αυτοματοποιημένα εργαλεία για την αξιολόγηση την νοητικής έκπτωσης σε πολύ πρώιμο στάδιο. Πιο συγκεκριμένα, στο σενάριο ταξινόμησης «Υγιής ομάδα» έναντι «Ομάδα με MCI» το EER (Equal Error Rate), δηλαδή το σημείο στο οποίο το ποσοστό του σφάλματος τύπου I ισούται με το ποσοστό του σφάλματος τύπου II, ήταν  $20\% \pm 5\%$ . Για το σενάριο «Υγιής Ομάδα» έναντι «Ομάδα με AD», η τιμή του EER που υπολογίστηκε ήταν  $13\% \pm 3\%$ , ενώ στην τελευταία μελέτη ταξινόμησης («Ομάδα με AD» έναντι «Ομάδα με MCI») το EER βρέθηκε  $19\% \pm 5\%$ .<sup>[10]</sup>

Οι Tóth et al. (2018) επικεντρώθηκαν στην διάκριση των ομιλητών με Ήπια Γνωστική Διαταραχή από τους υγιείς ομιλητές και αξιολόγησαν την ακρίβεια της δυαδικής ταξινόμησης χρησιμοποιώντας το εργαλείο ανοικτού κώδικα Weka και τους κατηγοριοποιητές Naive Bayes, SVM και Random Forests. Σε πρώτο στάδιο, χρησιμοποιήθηκαν όλες οι γλωσσικές και ακουστικές παράμετροι που καταγράφηκαν και αξιολογήθηκε η επίδοση των τριών κατηγοριοποιητών σε δύο περιπτώσεις: (1) η εξαγωγή των δεδομένων ομιλίας έγινε με μη αυτόματο τρόπο και (2) η εξαγωγή των δεδομένων ομιλίας έγινε με τη χρήση τεχνικών ASR. Ο κατηγοριοποιητής Naive Bayes είχε τη χειρότερη απόδοση και για τα δύο σύνολα χαρακτηριστικών, ενώ η καλύτερη ακρίβεια ( $71,4\%$ ) επιτεύχθηκε με το

Random Forests για το δεύτερο σύνολο δεδομένων και με το SVM για το πρώτο σύνολο. Οι τιμές  $F_1$  ήταν 76% και 75%, αντίστοιχα. Σε δεύτερο στάδιο, μελετήθηκε η επίδοση των παραπάνω κατηγοριοποιητών χρησιμοποιώντας μόνο τα στατιστικά σημαντικά χαρακτηριστικά. Στην περίπτωση του ταξινομητή SVM, η επίδοση ήταν χειρότερη σε σχέση με αυτή που επιτεύχθηκε με τη χρήση ολόκληρου του συνόλου των δεδομένων, ενώ η επίδοση του ταξινομητή Random Forests βελτιώθηκε ελαφρώς. Η καλύτερη τιμή της ακρίβειας ήταν 75% και επιτεύχθηκε για το σύνολο των δεδομένων που εξήχθησαν μέσω του εργαλείου ASR. Η αντίστοιχη τιμή  $F_1$  ήταν 78,8%.<sup>[9]</sup>

Οι Chlasta & Wołk (2021) χρησιμοποίησαν ακουστικά δεδομένα που συλλέχθηκαν μέσω της άσκησης της περιγραφής εικόνας (Η δοκιμασία της κλοπής μπισκότου) από υγιείς ομιλητές και ασθενείς με AD στο πλαίσιο της δοκιμασίας «2020 ADReSS Challenge» και σύγκριναν την επίδοση δύο συστημάτων για αυτόματη ανίχνευση της AD. Σε πρώτη προσέγγιση, χρησιμοποιήθηκε το μοντέλο VGGish (ένα προ-εκπαιδευμένο μοντέλο) για την αυτόματη εξαγωγή των ακουστικών παραμέτρων και η μέθοδος της Ανάλυσης Κύριου Συστατικού (Principal Component Analysis, PCA) για τη μείωση των αρχικών μεταβλητών και πραγματοποιήθηκαν πειράματα ταξινόμησης με συνήθεις κατηγοριοποιητές, όπως Μηχανή Διανυσμάτων Υποστήριξης (SVM), Γραμμική SVM, Ταξινομητής Perceptron και Ταξινομητής Πλησιέστερου Γείτονα (1NN για KNN με  $K=1$ ). Η συγκεκριμένη προσέγγιση πέτυχε μέγιστη τιμή ακρίβειας 59,1%. Σε δεύτερο στάδιο, παρουσιάστηκε το μοντέλο DemCNN, ένα ειδικό μοντέλο Συνελικτικών Νευρωνικών Δικτύων (Convolutional Neural Networks, CNNs), για την ανίχνευση ενδείξεων της νόσου Alzheimer στον προφορικό λόγο. Σύμφωνα με τα πειράματα, η προτεινόμενη αρχιτεκτονική πέτυχε μέγιστη τιμή ακρίβειας 63,6% και επέδειξε καλές δυνατότητες γενίκευσης.<sup>[2]</sup>

## 5 Συμπεράσματα

Η νόσος Alzheimer υποβαθμίζει, λοιπόν, τις γνωστικές και λεκτικές δεξιότητες των ασθενών, καθιστώντας δύσκολη την επικοινωνία. Ο λόγος, ως πρωτεύον μέσο της ανθρώπινης επικοινωνίας, παρέχει πληθώρα γλωσσικών και ακουστικών δεδομένων που είναι δυνατόν να χρησιμοποιηθούν στην έρευνα για την πρόωπη διάγνωση της AD. Ακόμη, η εμφάνιση διαταραχών του λόγου σε όλα τα στάδια της AD, καθώς και η άμεση διασύνδεση του λόγου και της γλώσσας με άλλες νοητικές και εκτελεστικές λειτουργίες<sup>[3],[7]</sup>, καθιστά τη μελέτη της ομιλίας αξιόπιστο και καθολικό μέτρο τόσο για την έγκαιρη διάγνωση της AD όσο και για την παρακολούθηση της εξέλιξης της πάθησης σε ασθενείς με AD. Με βάση τις πρόσφατες έρευνες<sup>[3],[9],[10]</sup>, υπάρχουν ενδείξεις πως αλλαγές στην ομιλία και τη γλώσσα μπορούν να ανιχνευτούν ακόμη και από το στάδιο της MCI· τα στοιχεία όμως για την διακριτική ικανότητα των χαρακτηριστικών του λόγου στο προκλινικό στάδιο της AD είναι περιορισμένα. Κλινικά, η μελέτη της ομιλίας μπορεί να αποτελέσει ένα γρήγορο και αποτελεσματικό μέσο παροχής δεδομένων, με χαμηλή επιβάρυνση για τους ασθενείς. Ακόμη, οι συνεχείς τεχνολογικές εξελίξεις στον τομέα της μηχανικής μάθησης, της αυτόματης αναγνώρισης ομιλίας και της επεξεργασίας φυσικής γλώσσας προσφέρουν πιο αποτελεσματικά και πρόσφορα μέσα για την ανάλυση του λόγου και των χαρακτηριστικών της γλώσσας.



Στις έρευνες που μελετήθηκαν, αναφέρονται υψηλές τιμές στην ακρίβεια των μοντέλων μηχανικής μάθησης για τη διάκριση ασθενών με AD, MCI και υγιών ατόμων, παρά τις εγγενείς δυσκολίες που απορρέουν από την ετερογένεια των δεδομένων και των μεθόδων. Λόγω της ανομοιογένειας των μεθόδων συλλογής δεδομένων (ποικίλες γνωστικές ασκήσεις για καταγραφή δεδομένων μονολόγου, διαλόγου, αυθόρμητης ή μη-αυθόρμητης ομιλίας), των γλωσσικών δομών και των ακουστικών χαρακτηριστικών που μελετώνται και των τεχνικών εξαγωγής και επεξεργασίας των δεδομένων, καθίσταται δύσκολη η ισότιμη σύγκριση των αποτελεσμάτων. Συνεπώς, η καθιέρωση μίας τυπικής ταξινόμησης των γλωσσικών και ακουστικών χαρακτηριστικών, όπως αυτή που πρότειναν οι Voletti et al. (2020) και ακολουθήθηκε στην παρούσα εργασία, κρίνεται απαραίτητη για την αντικειμενική σύγκριση της επίδοσης των διάφορων μοντέλων. Επιπλέον αδυναμίες που εντοπίζονται στις έρευνες είναι ο μικρός αριθμός δειγμάτων (συμμετεχόντων στα πειράματα) και η έλλειψη εθνικής και πολιτισμικής ποικιλομορφίας στα δεδομένα που συλλέγονται.

Προτείνουμε λοιπόν, η μελλοντική έρευνα να επικεντρωθεί σε δεδομένα φυσικής ομιλίας, τα οποία είναι δυνατό να καταγράφονται συνεχώς με παθητικό τρόπο (για παράδειγμα, εξαγωγή δεδομένων από μία συζήτηση μέσω του κινητού τηλεφώνου) και στην ανάπτυξη αυτοματοποιημένων μεθόδων επεξεργασίας και ανάλυσης των γλωσσικών δεδομένων, προκειμένου να ελαχιστοποιηθεί η μεροληψία. Η μελλοντική έρευνα κρίνεται ακόμη απαραίτητο να συμπεριλάβει, στις πειραματικές δοκιμές, ομάδες υγιών ατόμων με μη εμφανή συμπτώματα διαταραχής στον λόγο, αλλά υψηλό κίνδυνο εμφάνισης της AD, προκειμένου να εκτιμηθεί η ακρίβεια της διάγνωσης της AD σε πρώιμο, προκλινικό στάδιο από χαρακτηριστικά του λόγου και η χρήση δειγμάτων με εθνική και πολιτισμική ποικιλομορφία, για τον εντοπισμό πιθανών διαφορών. Σε αυτούς τους τομείς μπορεί να συμβάλει καταλυτικά η Τεχνητή Νοημοσύνη και το Διαδίκτυο των Πραγμάτων (Internet of Things), δεδομένης της συνεχώς αυξανόμενης παρουσίας αυτών των τεχνολογιών στην καθημερινή ζωή των ανθρώπων, μέσω των «έξυπνων» κινητών τηλεφώνων (smartphones), των φορητών συσκευών (wearables) και των «έξυπνων» σπιτιών (smart homes). Ένα ουσιαστικό πλεονέκτημα της χρήσης τέτοιων τεχνολογιών είναι η δυνατότητα δημιουργίας ενός μη-επεμβατικού, φιλικού προς τον χρήστη και οικονομικού ψηφιακού εργαλείου εξ αποστάσεως αξιολόγησης της γνωστικής υγείας των χρηστών, που θα παρέχει άμεση ανατροφοδότηση. Αυτή η συνεισφορά θεωρούμε πως είναι ιδιαίτερα σημαντική στην σημερινή εποχή με την πανδημία COVID-19, καθώς η ανάγκη εξ αποστάσεως παρακολούθησης της υγείας των ηλικιωμένων ατόμων και των ευπαθών ομάδων είναι μεγάλη.

## Βιβλιογραφικές Αναφορές

- [1] Burns A, Iliffe S. Alzheimer's disease BMJ 2009; 338 :b158 doi:10.1136/bmj.b158
- [2] Chlasta, K., Wołk, K. (2021). Towards Computer-Based Automated Screening of Dementia Through Spontaneous Speech. *Frontiers in psychology*, 11, 623237. <https://doi.org/10.3389/fpsyg.2020.623237>
- [3] Mueller, K. D., Hermann, B., Mecollari, J., & Turkstra, L. S. (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *Journal of clinical and experimental neuropsychology*, 40(9), 917–939. <https://doi.org/10.1080/13803395.2018.1446513>
- [4] Alzheimer's Association. 2021 Alzheimer's Disease Facts and Figures Alzheimer's and Dementia. Available from: <https://www.alz.org/media/Documents/alzheimers-facts-and-figures.pdf>
- [5] Alzheimer's Association. Causes and Risk Factors for Alzheimer's Disease. Available from: <https://www.alz.org/alzheimers-dementia/what-is-alzheimers/causes-and-risk-factors>
- [6] National Institute on Aging. What Happens to the Brain in Alzheimer's Disease? Available from: <https://www.nia.nih.gov/health/what-happens-brain-alzheimers-disease>
- [7] de la Fuente Garcia, Sofia, Ritchie, Craig W., and Luz, Saturnino. 'Artificial Intelligence, Speech, and Language Processing Approaches to Monitoring Alzheimer's Disease: A Systematic Review'. 1 Jan. 2020:1547 – 1574.
- [8] Ferris, S. H., Farlow, M. (2013). Language impairment in Alzheimer's disease and benefits of acetylcholinesterase inhibitors. *Clinical interventions in aging*, 8, 1007–1014. <https://doi.org/10.2147/CIA.S39959>
- [9] Toth, L., Hoffmann, I., Gosztolya, G., Vincze, V., Szatloczki, G., Banreti, Z., Pakaski, M., & Kalman, J. (2018). A Speech Recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech. *Current Alzheimer research*, 15(2), 130–138. <https://doi.org/10.2174/1567205014666171121114930>
- [10] König, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., Manera, V., Verhey, F., Aalten, P., Robert, P. H., & David, R. (2015). Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & dementia (Amsterdam, Netherlands)*, 1(1), 112–124. <https://doi.org/10.1016/j.dadm.2014.11.012>
- [11] Mah Parsa, Muhammad Raisul Alam, Alex Mihailidis et al. Towards AI-powered Language Assessment Tools, 25 February 2021, PREPRINT (Version 1) available at Research Square, <https://doi.org/10.21203/rs.3.rs-246079/v1>
- [12] R. Voleti, J. M. Liss and V. Berisha, "A Review of Automated Speech and Language Features for Assessment of Cognitive and Thought Disorders," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 2, pp. 282–298, Feb. 2020, doi: 10.1109/JSTSP.2019.2952087