



Καινοτόμες λύσεις για τη βιώσιμη και περιβαλλοντικά φιλική φυτοπροστασία των οπωροκηπευτικών της Ελλάδας, στην Ευρώπη του μέλλοντος

Παραδοτέο Π.1.5.1: Φασματικές τεχνολογίες για διάγνωση των ασθενειών ωίδιο σε αμπέλια και τεφρά σήψη σε τομάτα

Πληροφορίες για το έγγραφο

Αριθμός παραδοτέου: **Π.1.5.1**

Ενότητα εργασίας: **ΕΕ1**

Επικεφαλής δικαιούχος: **[ΑΠΘ]**

Συγγραφείς: **[Ξανθούλα Ειρήνη Πανταζή, Αντώνιος Μορέλλος, Χρήστος Τσιτσόπουλος]**

Έκδοση: **4.0**

Είδος Παραδοτέου: **[Έκθεση]**

Ημερομηνία παράδοσης: **[15 - 01 – 2025]**

Στοιχεία Πράξης

Τίτλος: Καινοτόμες λύσεις για τη βιώσιμη και περιβαλλοντικά φιλική φυτοπροστασία των οπωροκηπευτικών της Ελλάδας, στην Ευρώπη του μέλλοντος

Τίτλος (EN): InnoPP-Innovations in Plant Protection for sustainable and environmentally friendly pest control

Κωδικός πράξης: TAEDR-0535675

Ακρωνύμιο έργου: InnoPP

Ημερομηνία έναρξης: 15 Μαΐου 2023

Διάρκεια: 28 Μήνες

Συντονιστής Φορέας: Γεωπονικό Πανεπιστήμιο Αθηνών

Συντονιστής/ Επιστημονικός Υπεύθυνος: Ιωάννης Βόντας

Πίνακας Περιεχομένων

1	ΕΙΣΑΓΩΓΗ ΚΑΙ ΣΤΟΧΟΙ	4
2	ΠΕΡΙΓΡΑΦΗ ΤΩΝ ΕΡΓΑΣΙΩΝ.....	9
2.1	Υλικά και μεθοδοι.....	9
2.2	Αποτελέσματα και συζήτηση	20
3	ΣΥΝΟΨΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ ... Σφάλμα! Δεν έχει οριστεί σελιδοδείκτης.	51
4	ΠΑΡΑΡΤΗΜΑ Ι- Βιβλιογραφικές Αναφορές	52

Περίληψη του Έργου

Το έργο «Καινοτόμες λύσεις για τη βιώσιμη και περιβαλλοντικά φιλική φυτοπροστασία των οπωροκηπευτικών της Ελλάδας, στην Ευρώπη του μέλλοντος» στοχεύει στην ανάπτυξη σύγχρονων και καινοτόμων μεθόδων για την προστασία των καλλιεργειών όπως τα κηπευτικά, τα εσπεριδοειδή και το επιτραπέζιο σταφύλι. Περιλαμβάνει τη δημιουργία προηγμένων διαγνωστικών εργαλείων για την ανίχνευση εχθρών και παθογόνων με τεχνολογίες αιχμής, όπως ηλεκτρονικές παγίδες και βιοαισθητήρες, καθώς και πλατφόρμες αλληλούχισης για τον πλήρη προσδιορισμό των ιωμάτων. Επιπλέον, θα αναπτυχθούν μοντέλα πρόβλεψης επιδημιών και καινοτόμα βιοφυτοπροστατευτικά προϊόντα, τα οποία θα αξιολογηθούν για την ασφάλεια τους σε μη στόχους οργανισμούς. Τέλος, οι νέες τεχνολογίες θα ενσωματωθούν σε συστήματα ολοκληρωμένης διαχείρισης φυτοπροστασίας και θα δοκιμαστούν σε πραγματικές συνθήκες, ενώ θα αξιολογηθούν οι κοινωνικοοικονομικές και περιβαλλοντικές επιπτώσεις τους.

Σύνοψη της ΕΕ1

Η ΕΕ1 έχει ως στόχο την ανάπτυξη διαγνωστικών εργαλείων και μεθόδων για την ανίχνευση, ταυτοποίηση και παρακολούθηση εχθρών και παθογόνων, περιλαμβάνοντας καινοτόμες προσεγγίσεις όπως βελτιωμένα συστήματα παρακολούθησης εντομολογικών εχθρών με σύγχρονες παγίδες και νέα ελκυστικά. Στη πλαίσια υλοποίησης της ΕΕ θα αναπτυχθούν σύγχρονα μοριακά διαγνωστικά εργαλεία, όπως η τεχνολογία LAMP και kit «Ready to Go Lyophilised pellets» για ποιοτική και ποσοτική ταυτοποίηση, καθώς και υπερευαίσθητες πλατφόρμες ακριβείας. Επιπλέον, θα χρησιμοποιηθούν τεχνολογίες τεχνητής νοημοσύνης και φασματικές τεχνικές για την αναγνώριση ζιζανίων και ασθενειών, ενώ βιοαισθητήρες θα χρησιμοποιηθούν για την ανίχνευση υπολειμμάτων φυτοφαρμάκων και ωχρατοξινών. Τέλος, θα αναπτυχθούν νέοι μοριακοί δείκτες ανθεκτικότητας και διαγνωστικά συστήματα για την αξιολόγηση της άμυνας των φυτών και της βιοδραστικότητάς τους.

Συνοπτική παρουσίαση του παραδοτέου Π.1.5.1

Ο σκοπός του παρόντος παραδοτέου είναι η ανάπτυξη φασματικών τεχνολογιών για τη διάγνωση των ασθενειών ωίδιο σε αμπέλια και τεφρά σήψη σε τομάτες. Η έρευνα που διεξάχθηκε στα πλαίσια υλοποίησης του παραδοτέου αυτού, εστιάζει στον εντοπισμό των ασθενειών σε προσυμπτωματικά στάδια μέσω τεχνητής νοημοσύνης και υπερφασματικής ανάλυσης, με στόχο τη μείωση της χρήσης χημικών προϊόντων και την ενίσχυση της βιώσιμης γεωργίας.

Οι εργασίες οι οποίες πραγματοποιήθηκαν στα πλαίσια υλοποίησης του παρόντος παραδοτέου περιελάμβαναν: 1. την ανάπτυξη και εφαρμογή πολύπλοκων τεχνικών για τη συλλογή και ανάλυση φασματικών δεδομένων από μολυσμένα και υγιή φυτά καλλιέργειας αμπελιού και τομάτας, 2. την ανάπτυξη διαφορετικών μοντέλων τεχνητής νοημοσύνης για την ταξινόμηση των φασματικών δεδομένων και την εξαγωγή πληροφοριών για τη φαινοτυπική κατάσταση των φυτών και την ανάπτυξη πειραματικών πρωτοκόλλων και διαδικασιών προεπεξεργασίας δεδομένων ικανών να

εξασφαλίσουν μεγάλη ακρίβεια.

Δημιουργήθηκαν 2 πρότυπα ανίχνευσης για τα παθογόνα ωίδιο και τεφρά σήψη σε καλλιέργεια αμπελιού και τομάτας αντίστοιχα. Όσο αφορά την ανίχνευση του ωιδίου στον αμπελώνα, αναπτύχθηκαν τρία μοντέλα μηχανικής μάθησης και συγκεκριμένα το μοντέλο Random Forest (RF), K-Nearest Neighbors (KNN) και Multilayer Perceptron (MLP). Όλα τα μοντέλα παρουσίασαν υψηλή ακρίβεια όσο αφορά την ανίχνευση του παθογόνου σε προσυμπτωματικό στάδιο (accuracy = 98-99%), με το μοντέλο Random Forest να επιτυγχάνει ακρίβεια εκπαίδευσης και ελέγχου 99%.

Για το πείραμα που πραγματοποιήθηκε στην καλλιέργεια τομάτας, αναπτύχθηκαν τα ίδια μοντέλα μηχανικής μάθησης (Random Forest, KNN και MLP) για την ανίχνευση της τεφράς σήψης. Τα μοντέλα εκπαιδεύτηκαν και αξιολογήθηκαν για την ταξινόμηση των φυτών σε τρεις κατηγορίες: υγιή, μολυσμένα χωρίς ορατά συμπτώματα και μολυσμένα με ορατά συμπτώματα. Τα δεδομένα προ-επεξεργάστηκαν για τη βελτίωση της απόδοσης των μοντέλων και περιλάμβαναν 8 κατηγορίες βάσει του χρόνου μετά τη μόλυνση (0h, 4h, 8h, 12h, 24h, 48h, 72h, 96h), με το μοντέλο MLP να αποδίδει καλύτερα στη διάκριση των μετέπειτα φάσεων της μόλυνσης με ακρίβεια (accuracy) 81% και F1-Score ίσο με 81% στη φάση ελέγχου.

Οι τεχνικές που εφαρμόστηκαν και παρουσιάζονται στο παρόν παραδοτέο, ανέδειξαν την αξιοπιστία και την αποτελεσματικότητα που προσφέρει ο συνδυασμός της φασματικής ανάλυσης με τη χρήση τεχνητής νοημοσύνης, όσο αφορά την έγκαιρη ανίχνευση φυτοπαθολογικών ασθενειών. Η προσέγγιση αυτή αποδεικνύεται ικανή να παρέχει ακριβή και έγκαιρη διάγνωση, ενισχύοντας τη βιωσιμότητα των γεωργικών πρακτικών και μειώνοντας την εξάρτηση από παραδοσιακές και μη στοχευμένες τεχνικές εφαρμογής φυτοπροστατευτικών σκευασμάτων .

1 ΕΙΣΑΓΩΓΗ ΚΑΙ ΣΤΟΧΟΙ

Η γεωργία διαδραματίζει σημαντικό ρόλο στην παγκόσμια οικονομική δραστηριότητα. Η συνεχιζόμενη αύξηση του παγκόσμιου πληθυσμού έχει συνέπεια την μεγαλύτερη πίεση στον τομέα της γεωργίας. Νέες μέθοδοι αγρό-τεχνολογιών αναπτύσσονται στον τομέα της Γεωργίας Ακριβείας με σκοπό την βελτιστοποίηση της παραγωγικότητας (Liakos et al., 2018).

Οι ασθένειες των φυτών αποτελούν σημαντική απειλή για την γεωργία προκαλώντας απώλειες της τάξεως του 10-16% της παγκόσμιας παραγωγής (Golhani et al., 2018). Σε συνδυασμό με την προβλεπόμενη αύξηση του παγκοσμίου πληθυσμού σε 9 δισεκατομμύρια μέχρι το έτος 2050 σύμφωνα με τα στοιχεία του Food and Agriculture Organization (FAO), η γεωργική παραγωγή πρέπει να παρουσιάσει αύξηση της τάξεως του 70% για να καλυφθούν οι ανάγκες του πληθυσμού. Συνεπώς, ο γεωργικός τομέας στρέφεται στην στρατηγική της Γεωργίας Ακριβείας ώστε να υπάρξει αύξηση της παραγωγής από τις υπάρχουσες καλλιεργούμενες εκτάσεις, μειώνοντας παράλληλα τις περιβαλλοντικές επιπτώσεις, με την χρήση και ενσωμάτωση νέων τεχνολογιών στην γεωργία (Behmann et al., 2015).

Η συνήθης πρακτική αντιμετώπισης ασθενειών είναι ο ψεκασμός των φαρμάκων στην περιοχή του αγρού ή του θερμοκηπίου. Παρότι η αντιμετώπιση των ασθενειών με αυτόν τον τρόπο είναι αποτελεσματική, το οικονομικό κόστος είναι μεγάλο αλλά το κυριότερο είναι ότι δημιουργούνται ζητήματα περιβαλλοντικής φύσεως λόγω της μαζικής χρήσης χημικών ουσιών και της επίδρασης που μπορεί να έχουν με την παρουσία υπολειμμάτων στο τελικό προϊόν ή της ρύπανσης των υδάτων της περιοχής (Liakos et al., 2018). Έτσι,

δημιουργείται η ανάγκη για τη στοχευμένη διάγνωση και αντιμετώπιση των ασθενειών.

Παραδοσιακά, η διάγνωση των ασθενειών στα φυτά περιλαμβάνει τεχνικές όπως η οπτική αναγνώριση των συμπτωμάτων ή και την καταστροφή του φυτού ώστε μετέπειτα να πραγματοποιηθεί χημική ανάλυση τους σε συνθήκες εργαστηρίου (Strange and Scott, 2005). Η διάγνωση των ασθενειών που προκαλούνται από μύκητες συνήθως πραγματοποιείται με βάση τα μορφολογικά χαρακτηριστικά έπειτα από απομόνωση του προσβεβλημένου φυτού ενώ οι ιολογικές ασθένειες ταυτοποιούνται με εξαγωγή και ανάλυση του γενετικού υλικού του ιού. Αυτές οι τεχνικές βασίζονται στη δειγματοληψία φύλλων και την μετέπειτα χημική ανάλυση τους, γεγονός που τις χαρακτηρίζει ως τεχνικές καταστρεπτικές ως προς το φυτό (Golhani et al., 2018).

Κατά τον 21ο αιώνα αναπτύσσονται συνεχώς νέες τεχνικές με τη χρήση αυτοματοποιημένων μεθόδων για την αναγνώριση και ταυτοποίηση προβλημάτων μέσω αυτοματοποιημένων τεχνικών υπολογιστικής νοημοσύνης. Η πρόοδος της τεχνολογίας σε επίπεδο τόσο της υπολογιστικής δύναμης των υπολογιστών επιτρέπει την αξιοποίηση τεχνικών όπως τα τεχνητά νευρωνικά δίκτυα για ανάλυση μεγάλου όγκου δεδομένων και αναγνώριση και πρόβλεψη μοτίβων και προβλημάτων. Τεχνικές αυτού του τύπου αξιοποιούνται πλέον από ένα ευρύ φάσμα επιστημών για την εξαγωγή συμπερασμάτων.

Ο γεωργικός τομέας παρουσιάζει, αντίστοιχα, αύξηση της χρήσης μεθόδων υπολογιστικής νοημοσύνης για την ανάλυση προβλημάτων, ειδικών στις γεωργικές επιστήμες. Μοντέλα έχουν αξιολογηθεί για την πρόβλεψη της εξατμισοδιαπνοής (Feng et al., 2017; Patil and Deka, 2016), στην εκτίμηση της εδαφικής θερμοκρασίας (Nahvi et al., 2016), ενώ οι Morellos et al. (2016) έδειξαν την δυνατότητα πρόβλεψης της εδαφικής υγρασίας, αζώτου και οργανικού άνθρακα με χρήση της μεθόδου LS-SVM. Παράλληλα, παρουσιάζονται μελέτες στον τομέα του εντοπισμού ζιζανίων. Οι Karimi et al. (2006) μελέτησαν μοντέλο SVM για τον εντοπισμό ζιζανίου και καταπόνηση αζώτου σε καλαμπόκι. Οι LÓPEZ-GRANADOS et al. (2008) χρησιμοποίησαν πολυφασματικούς αισθητήρες για την αναγνώριση ζιζανίων στο σιτάρι ενώ παρουσιάστηκε μέθοδος εντοπισμού ζιζανίων σε καλλιέργεια σόγιας (Suzuki et al., 2008). Οι Tamouridou et al. (2018) παρουσίασαν χαρτογράφηση ζιζανίου με χρήση UAV.

Ο εντοπισμός ασθενειών αποτελεί μία από τις βασικότερες ανησυχίες του γεωργικού τομέα, είτε σε συνθήκες θερμοκηπίου, είτε ανοικτού αγρού (Liakos et al., 2018). Παραδοσιακά, η ταυτοποίηση ενός μύκητα πραγματοποιείται μορφολογικά και έπειτα με απομόνωση και καλλιέργεια του ενώ οι ιοί ταυτοποιούνται με βάση το γενετικό τους υλικό. Οι παραδοσιακές τεχνικές μπορεί να στερούνται ακρίβειας και συνέπειας αποτελεσμάτων και να απαιτούν σημαντικό κόστος (Sakudo et al., 2006; Thanarajoo et al., 2014).

Τα πλεονεκτήματα που εμφανίζουν τεχνικές που βασίζονται σε φασματοσκοπικές μεθόδους, έναντι των παραδοσιακών παρεμβατικών μεθόδων, για την εξαγωγή των δεδομένων για την ταυτοποίηση των ασθενειών, σε συνδυασμό με τις τεχνικές της μηχανικής μάθησης είναι ότι αποτελούν μη επεμβατικές τακτικές προς το φυτό, χωρίς να απαιτούν την καταστροφή του, πραγματοποιούνται στον χώρο όπου καλλιεργείται το φυτό και μπορούν να επιφέρουν μικρότερο κόστος σε σχέση με τις παραδοσιακές

εργαστηριακές αναλύσεις και αποτελέσματα σε συντομότερο χρονικό διάστημα. Περιπτώσεις τέτοιων μελετών παρουσιάζονται στην βιβλιογραφία ολοένα και συχνότερα τα τελευταία έτη. Ήδη από το 2004 οι Moshou et al. (2004) ανίχνευσαν, με τη χρήση υπερφασματικής απεικόνισης και τεχνητών νευρωνικών δικτύων, την κίτρινη σκωρίαση του σιταριού με χρήση χαρτών Kohonen για την ανίχνευση ασθενειών με βάση υπερφασματικά και πολυφασματικά δεδομένα (Moshou et al., 2005). Οι Wang et al. (2008) παρουσίασαν μελέτη για τον εντοπισμό μόλυνσης από *Phytophthora* σε φυτά ντομάτας με χρήση νευρωνικών δικτύων. Έρευνες επίσης σχετικά με τη διάκριση καταπόνησης από βιοτικούς έναντι αβιοτικούς παράγοντες έχουν πραγματοποιηθεί με χρήση δικτύων Αυτό-οργανωμένων Χαρτών (Self-Organizing Maps, SOM) (Pantazi et al., 2017a) και Μηχανών Διανυσματικής Υποστήριξης των Ελαχίστων Τετραγώνων (Least Square Support Vector Machines, LS-SVMs) (Moshou et al., 2014). Δίκτυα Μηχανών Διανυσματικής Υποστήριξης (Support, Vector Machines, SVM) έχουν χρησιμοποιηθεί για τον εντοπισμό παρασίτων σε φυτά φράουλας (Ebrahimi et al., 2017) και της ασθένειας *Bakanae* σε καλλιέργειες ρυζιού (Chung et al., 2016). Οι Pantazi et al. (2017b) παρουσίασαν την δυνατότητα επιτυχούς αναγνώρισης φυτών *Silybum Marianum* που έχουν προσβληθεί από τον μύκητα *Microbotryum Silybum*, ενώ ο (Ferentinos, 2018) παρουσίασε μια γενικευμένη μέθοδο ταξινόμησης φυτών σε υγιή και ασθενή με βάση εικόνες RGB των φύλλων τους και ένα Συνελικτικό (convolutional) νευρωνικό δίκτυο.

Η ασθένεια της τεφράς σήψης σε φυτά τομάτας

Ο μύκητας *Botrytis cinerea*, μέλος του γένους *Botrytis* και της οικογένειας *Sclerotiniaceae*, αποτελεί έναν από τους πιο σημαντικούς παθογόνους μικροοργανισμούς που προσβάλλουν καλλιέργειες τομάτας. Ο μύκητας αυτός ευθύνεται για την ασθένεια γνωστή ως "τεφρά σήψη" και αποτελεί παγκόσμια απειλή για πολλές καλλιέργειες, καθώς είναι εξαιρετικά προσαρμοστικός σε διάφορες κλιματικές συνθήκες (Williamson et al., 2007). Το *Botrytis cinerea* εντοπίστηκε για πρώτη φορά τον 19ο αιώνα και έκτοτε αποτελεί σημαντικό αντικείμενο μελέτης λόγω της ευρείας εξάπλωσής του και των οικονομικών επιπτώσεων που δύναται να προκαλέσει (Dean et al., 2012).

Ο *Botrytis cinerea* μεταδίδεται κυρίως μέσω σπορίων που μεταφέρονται από τον άνεμο και το νερό. Τα σπόρια αυτά μπορούν να παραμείνουν ζωντανά στο έδαφος ή στα υπολείμματα των καλλιεργειών για μεγάλες χρονικές περιόδους, καθιστώντας τον έλεγχο εξέλιξης και εξάπλωσης του μύκητα ιδιαίτερα δύσκολο (Fillinger & Elad, 2016). Επιπλέον, η μετάδοση της προσβολής μπορεί να ενισχυθεί μέσω μηχανικής μεταφοράς, όπως για παράδειγμα μέσω της ανθρώπινης δραστηριότητας κατά τη συγκομιδή ή τις καλλιεργητικές φροντίδες/δαστηριότητες (Fillinger & Elad, 2016).

Τα συμπτώματα της τεφράς σήψης σε φυτά τομάτας περιλαμβάνουν την εμφάνιση υδατώδους κηλίδων στους καρπούς, τα φύλλα και τους μίσχους. Αρχικά, οι μολυσμένες περιοχές μπορεί να εμφανίζουν μια υγρή όψη, η οποία γρήγορα εξελίσσεται σε καφέ, νεκρωτικές κηλίδες που καλύπτονται με χαρακτηριστικό γκριζωπό μύκητα. Οι μολύνσεις συνήθως ξεκινούν από τραυματισμένα ή εξασθενημένα μέρη των φυτών, όπως πληγές ή περιοχές με μηχανική καταπόνηση, που επιτρέπουν στον μύκητα να εισέλθει πιο εύκολα. Στους καρπούς, η προσβολή συχνά ξεκινάει από το σημείο επαφής με το

έδαφος ή με άλλες επιφάνειες, προκαλώντας μαλακή σήψη που εξελίσσεται γρήγορα (Jarvis, 1977). Καθώς η ασθένεια εξελίσσεται, οι καρποί γίνονται μη εμπορεύσιμοι λόγω της έντονης υποβάθμισης της ποιότητας και της εμφάνισής τους.

Επιπλέον, η προσβολή από τον *Botrytis cinerea* μπορεί να προκαλέσει φυλλόπτωση, καθώς τα φύλλα που έχουν προσβληθεί νεκρώνονται και πέφτουν. Αυτή η κατάσταση μειώνει τη φωτοσυνθετική ικανότητα του φυτού, επηρεάζοντας αρνητικά την ανάπτυξη και την παραγωγικότητά του (Elmer et al., 2007).

Στην Ελλάδα, ο *Botrytis cinerea* εντοπίζεται σε πολλές περιοχές όπου καλλιεργούνται τομάτες, ιδιαίτερα σε θερμοκήπια και περιοχές με αυξημένη υγρασία, όπως η Κεντρική Μακεδονία και η Κρήτη (Pappas, 2000). Οι ιδανικές συνθήκες για την ανάπτυξη του μύκητα περιλαμβάνουν υψηλά επίπεδα υγρασίας και θερμοκρασίες μεταξύ 15-25°C, γεγονός που καθιστά ορισμένες περιοχές της χώρας ιδιαίτερα ευάλωτες στις επιδημίες της ασθένειας.

Ο *Botrytis cinerea* ευδοκίμει σε συνθήκες υψηλής υγρασίας και μέτριων θερμοκρασιών. Η ύπαρξη ελεύθερου νερού στην επιφάνεια των φυτών είναι κρίσιμη για την επιβίωση και ανάπτυξη του μύκητα, καθώς το νερό διευκολύνει τη βλάστηση των σπορίων και τη μόλυνση του φυτού (Jarvis, 1980). Οι βροχοπτώσεις και η δροσιά αποτελούν σημαντικούς παράγοντες για την ανάπτυξη της ασθένειας, ενώ οι πυκνές φυτεύσεις και η κακή αερισμότητα επιτείνουν την εξάπλωση του μύκητα (Elad et al., 2007).

Η αντιμετώπιση του *Botrytis cinerea* σε προσυμπτωματικό επίπεδο αποτελεί σημαντική πρόκληση για την επιστημονική κοινότητα. Πρόσφατες έρευνες επικεντρώνονται στην ανάπτυξη μεθόδων που επιτρέπουν την έγκαιρη ανίχνευση και την προληπτική αντιμετώπιση του μύκητα πριν την εμφάνιση ορατών συμπτωμάτων. Με τη χρήση τεχνολογιών φασματικής ανάλυσης και τεχνητής νοημοσύνης, οι ερευνητές έχουν καταφέρει να αναπτύξουν μοντέλα που μπορούν να ανιχνεύσουν μικρές αλλαγές στη φασματική υπογραφή των φυτών, οι οποίες σχετίζονται με την αρχική μόλυνση από τον μύκητα (Raza et al., 2020). Επιπλέον, η χρήση βιολογικών παραγόντων, όπως ωφέλιμων μικροοργανισμών και αντιμυκητιασικών εκχυλισμάτων, έχει δείξει πολλά υποσχόμενα αποτελέσματα στη μείωση της ανάπτυξης του μύκητα σε πρώιμο στάδιο (Oni et al., 2022).

Οι τεχνικές αυτές έχουν τη δυνατότητα να βελτιώσουν σημαντικά την αποτελεσματικότητα των προληπτικών μέτρων και να μειώσουν την ανάγκη για χημικά φυτοπροστατευτικά, συμβάλλοντας έτσι σε μια πιο βιώσιμη γεωργική πρακτική.

Η ασθένεια του ωιδίου σε φυτά Αμπέλου

Ο μύκητας *Erysiphe necator* (συνώνυμο: *Uncinula necator*), μέλος της οικογένειας *Erysiphaceae*, αποτελεί έναν από τους πιο σημαντικούς παθογόνους οργανισμούς που προσβάλλουν τα αμπέλια. Ο μύκητας αυτός ευθύνεται για την ασθένεια γνωστή ως "ωίδιο" και προκαλεί σοβαρές οικονομικές απώλειες σε αμπελουργικές εκμεταλλεύσεις παγκοσμίως (Gadoury et al., 2012). Το *Erysiphe necator* καταγράφηκε για πρώτη φορά στην Ευρώπη το 1845 και από τότε έχει αποτελέσει αντικείμενο εκτεταμένης μελέτης λόγω της ταχείας εξάπλωσής του και των επιπτώσεων που επιφέρει στην παραγωγή και

ποιότητα του κρασιού (Pearson & Gadoury, 1987).

Ο μύκητας μεταδίδεται κυρίως μέσω ασκοσπορίων και κονιδίων, τα οποία διασπείρονται με τον άνεμο. Οι συνθήκες υψηλής υγρασίας και οι μέτριες θερμοκρασίες (20-25°C) ευνοούν τη βλάστηση των σπορίων και τη μόλυνση των φυτών (Calonnec et al., 2004). Οι κύριες πηγές μόλυνσης περιλαμβάνουν τα μολυσμένα φυτικά υπολείμματα και τα λανθάνοντα μολύσματα στους οφθαλμούς, τα οποία ενεργοποιούνται την άνοιξη υπό ευνοϊκές συνθήκες (Gadoury et al., 2001).

Τα συμπτώματα του ωιδίου στα αμπέλια περιλαμβάνουν χαρακτηριστική λευκή έως γκριζωπή εξάνθηση στις επιφάνειες των φύλλων, των βλαστών και των σταφυλιών. Στα φύλλα, εμφανίζονται κιτρινωπές κηλίδες που στη συνέχεια καλύπτονται από μυκηλιακό ιστό, ενώ σε σοβαρές περιπτώσεις τα φύλλα παραμορφώνονται και νεκρώνονται (Wilcox et al., 2015). Στους βλαστούς, το ωίδιο προκαλεί επιβράδυνση της ανάπτυξης και καχεξία, ενώ στους καρπούς δημιουργούνται σκασίματα και φελλοειδείς κηλίδες που καθιστούν τα σταφύλια ακατάλληλα για οινοποίηση (Jarvis, 1994).

Η προσβολή των σταφυλιών από το ωίδιο μπορεί να μειώσει σημαντικά την ποιότητα και την ποσότητα της παραγωγής, ενώ μπορεί να επηρεάσει και τα οργανοληπτικά χαρακτηριστικά του κρασιού. Ιδιαίτερα, η μόλυνση κατά τη φάση της ωρίμανσης των σταφυλιών είναι κρίσιμη, καθώς μπορεί να οδηγήσει σε σημαντική απώλεια σακχάρων και αρωμάτων (Calonnec et al., 2006).

Στην Ελλάδα, το ωίδιο αποτελεί μία από τις κυριότερες ασθένειες των αμπελιών, ιδιαίτερα σε περιοχές με αυξημένη υγρασία και μέτριες θερμοκρασίες, όπως η Μακεδονία, η Πελοπόννησος και τα νησιά του Αιγαίου (Tzitzikas et al., 2008). Οι συνθήκες αυτές δημιουργούν ένα ιδανικό περιβάλλον για την ανάπτυξη και εξάπλωση του μύκητα.

Η καταπολέμηση του *Erysiphe necator* περιλαμβάνει κυρίως προληπτικά μέτρα και εφαρμογές μυκητοκτόνων. Η χρήση θείου, είτε σε σκόνη είτε σε υγρή μορφή, παραμένει μια από τις πιο διαδεδομένες και αποτελεσματικές μεθόδους ελέγχου (Gubler et al., 1996). Ωστόσο, η υπερβολική χρήση χημικών μπορεί να οδηγήσει σε ανάπτυξη ανθεκτικότητας του μύκητα, καθώς και σε περιβαλλοντική επιβάρυνση.

Πρόσφατες έρευνες εστιάζουν στην ανάπτυξη τεχνολογικών εργαλείων για την έγκαιρη ανίχνευση και παρακολούθηση της ασθένειας. Με τη χρήση φασματικής ανάλυσης και αλγορίθμων τεχνητής νοημοσύνης, οι επιστήμονες έχουν αναπτύξει συστήματα που μπορούν να ανιχνεύσουν τις πρώιμες αλλαγές στα φυτά, πριν την εμφάνιση ορατών συμπτωμάτων (Pimentel et al., 2020). Επιπλέον, η χρήση βιολογικών παραγόντων, όπως ωφέλιμων μυκήτων και βακτηρίων, έχει δείξει ενθαρρυντικά αποτελέσματα στη μείωση των λοιμώξεων (Gaforio et al., 2019).

Η ενσωμάτωση αυτών των τεχνολογιών και πρακτικών μπορεί να μειώσει την εξάρτηση από τα χημικά φυτοπροστατευτικά, προάγοντας παράλληλα μια πιο βιώσιμη και φιλική προς το περιβάλλον αμπελουργία.

Σκοπός του παρόντος παραδοτέου ο εντοπισμός των ασθενειών στο αρχικό στάδιο της ανάπτυξης του παθογόνου, πριν την εμφάνιση των ορατών συμπτωμάτων της ασθένειας, σε φυτά ντομάτας και αμπέλου με χρήση φασματικών τεχνολογιών και αλγορίθμων τεχνητής νοημοσύνης, συγκεκριμένα μηχανικής εκμάθησης. Πέρα από την ανίχνευση της πιθανής βιοτικής καταπόνησης στα φυτά από τα παθογόνα τα οποία εξετάστηκαν, επιπλέον δόθηκε ιδιαίτερη προσοχή και στην ανίχνευση της δριμύτητας της επίδρασης της ασθένειας στα φυτά. Αυτό επιτυγχάνεται μέσω της δημιουργίας διαφορετικών κλάσεων που αντιστοιχούν σε ένα επίπεδο σοβαρότητας της ασθένειας, ανάλογα με τις ημέρες που παρήλθαν από την προσβολή του φυτού από το υπό μελέτη παθογόνο. Ο διαχωρισμός πραγματοποιήθηκε με τη βοήθεια της ερευνητικής ομάδας του Εργαστηρίου Φυτοπαθολογίας του τμήματος Γεωπονίας του Α.Π.Θ, υπό την επίβλεψη του διευθυντή του εργαστηρίου κ. Καραογλανίδη Γεώργιου.

Ο σκοπός του παρόντος εγγράφου είναι η παρουσίαση των αποτελεσμάτων καινοτόμου έρευνας στα πλαίσια του παραδοτέου Π1.5.1. Γίνεται εκτενής περιγραφή των πειραμάτων και των τεχνικών και τεχνολογιών που χρησιμοποιήθηκαν για την ανίχνευση των καταπονήσεων, ενώ παράλληλα παρουσιάζονται και τα αποτελέσματα των πειραμάτων. Τέλος, παρουσιάζονται και σχολιάζονται εκτενώς τα βασικά συμπεράσματα που προέκυψαν από αυτή την έρευνα.

Το παρόν έγγραφο **ακολουθεί την παρακάτω δομή:**

- 1. Εισαγωγή και Στόχοι:** Παρουσιάζεται το πλαίσιο της έρευνας και οι στόχοι του εγγράφου.
- 2. Περιγραφή των Εργασιών:** : 2.1. Υλικά και Μέθοδοι, 2.2. Αποτελέσματα και Συζήτηση.
- 3. Σύνοψη και Συμπεράσματα:** Βασικά ευρήματα της έρευνας και σχετικά συμπεράσματα.
- 4. Παράρτημα:** Βιβλιογραφικές αναφορές

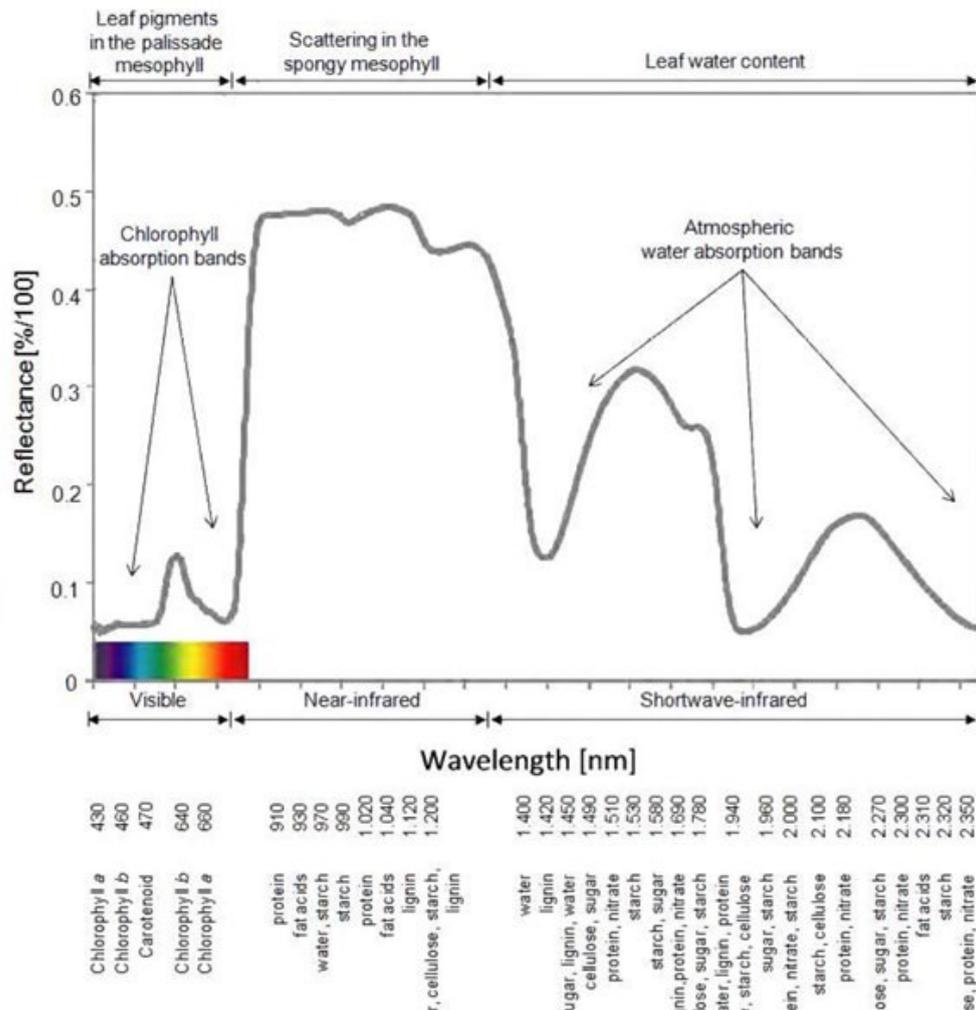
2 ΠΕΡΙΓΡΑΦΗ ΤΩΝ ΕΡΓΑΣΙΩΝ

2.1 Υλικά και Μέθοδοι Τεχνική Φασματοσκοπίας

Η φασματοσκοπία είναι η μελέτη της ηλεκτρομαγνητικής ακτινοβολίας και των αλληλεπιδράσεων της με την ύλη (Herrmann & Onkelinx, 1986). Η μελέτη της φασματοσκοπίας ξεκίνησε με την λήψη του φάσματος του ορατού φωτός καθώς διαπερνάει ένα πρίσμα και η διάθλαση των διαφορετικού μήκους κύματος ακτινοβολιών χωρίζει το φάσμα του ορατού από την ιώδη έως την ερυθρή ακτινοβολία. Στην συνέχεια, η έννοια της φασματοσκοπίας επεκτάθηκε στην ανάλυση όλου του φάσματος της ηλεκτρομαγνητικής ακτινοβολίας. Η αλληλεπίδραση της ύλης με την ακτινοβολία γίνεται από επίπεδο ατόμων και μορίων καθώς και των ιόντων τους ως και τα μεγαλύτερα σώματα. Οι βασικότεροι τρόποι αλληλεπίδρασης είναι η απορρόφηση της ενέργειας των φωτονίων της ακτινοβολίας από την ύλη, η μεταφορά της ακτινοβολίας μέσω των υλικών σωμάτων και η ανάκλαση της από τις επιφάνειες των σωμάτων και σωματιδίων (Baranoski & Rokne, 2001). Ανάλογα τον τύπο αυτών των αλληλεπιδράσεων, ορίζονται και τα αντίστοιχα πεδία της φασματοσκοπίας που σκοπό έχουν να μελετήσουν να εξαγάγουν συμπεράσματα από τον τρόπο που αντιδρά η ηλεκτρομαγνητική ακτινοβολία στις διάφορες καταστάσεις της ύλης.

Η χρήση της φασματοσκοπίας στις γεωπονικές επιστήμες παρουσιάζει ιδιαίτερο ενδιαφέρον στο επίπεδο των φύλλων των φυτικών οργανισμών. Τα φυτά βασίζονται στην διαδικασία της φωτοσύνθεσης για την μετατροπή της ηλιακής ενέργειας σε χημική. Πρωταρχικό στάδιο της φωτοσύνθεσης αποτελεί η απορρόφηση της ηλεκτρομαγνητικής ακτινοβολίας από την χλωροφύλλη στα φύλλα. Συνεπώς, η αλληλεπίδραση της ηλεκτρομαγνητικής ακτινοβολίας με αυτά αποτελεί σημαντικό πεδίο μελέτης (Jacquemoud & Ustin, 2001)

Το φάσμα του ορατού φωτός (VIS) εντοπίζεται από τα 400 έως τα 700nm. Σε αυτό το μέρος του φάσματος παρουσιάζονται υψηλά επίπεδα απορρόφησης της ακτινοβολίας λόγω της παρουσίας χρωστικών ουσιών. Συγκεκριμένα, η χλωροφύλλη α και η χλωροφύλλη β απορροφούν την ιώδη-κυανή (400 με 475nm) και την κίτρινη-ερυθρή ακτινοβολία (620 με 700nm), τα καροτενοειδή την ιώδη-κυανή (400 με 495nm) (Sims & Gamon, 2002) ενώ οι ανθοκυάνες απορροφούν την ακτινοβολία από την υπεριώδη έως και τα 550nm παρουσιάζοντας χαρακτηριστικά μέγιστα στα σημεία αυτά (Εικόνα 1.5.1-1). Ιδιαίτερο ενδιαφέρον παρουσιάζει η περιοχή του κοντινού υπέρυθρου (NIR) μεταξύ των 700 και 1100nm. Σε αυτήν την περιοχή η απορρόφηση είναι μικρότερη και εξαρτάται από τις στερεές ουσίες του φύλλου και την κυτταρική δομή (Mishra et al., 2017) (Bannari et al., 1995), συνεπώς η ανάκλαση είναι μεγαλύτερη. Η ανάκλαση της ακτινοβολίας σε αυτό το επίπεδο επηρεάζεται από τα χαρακτηριστικά της επιδερμίδας του φύλλου και από την διάχυση της ακτινοβολίας στην εσωτερική δομή του (Jensen, 2007).



Εικόνα 1.5.1-1. Φασματική ανάκλαση της βλάστησης και παράγοντες που επηρεάζουν την ανάκλαση των φύλλων στις περιοχές VIS, NIR, και SWIR και χαρακτηριστικά απορρόφησης βιοχημικών συστατικών των φυτών (Jensen, 2007)

Όσον αφορά στον τομέα του εντοπισμού ασθενειών, η υπόθεση που μελετάται είναι πως τα υγιή φυτά αλληλοεπιδρούν διαφορετικά με την ηλεκτρομαγνητική ακτινοβολία σε σχέση με ασθενή φυτά (D. Moshou et al., 2011). Κάποια συμπτώματα που αφορούν το πεδίο του ορατού φάσματος μπορούν να γίνουν αντιληπτά με το γυμνό μάτι αλλά με την χρήση εξειδικευμένου εξοπλισμού μπορούν να γίνουν αντιληπτά τα χαρακτηριστικά όλου του φάσματος.

Ταυτοποίηση Ασθενειών με χρήση Τεχνητής Νοημοσύνης

Η χρήση τεχνητής νοημοσύνης (AI) στον εντοπισμό και την ταυτοποίηση ασθενειών φυτών προσφέρει σημαντικά πλεονεκτήματα, όπως η ακρίβεια, η ταχύτητα και η μη

καταστροφική συλλογή δεδομένων. Με την αξιοποίηση υπερφασματικών τεχνικών, οι αλγόριθμοι ΑΙ μπορούν να ανιχνεύσουν διαφορές στη φασματική υπογραφή μεταξύ υγιών και μολυσμένων φυτών, ακόμη και σε προσυμπτωματικά στάδια. Αυτές οι διαφορές αντικατοπτρίζουν φυσιολογικές και βιοχημικές αλλαγές που προκαλούνται από παθογόνους μικροοργανισμούς. Συγκεκριμένα, αλγόριθμοι μηχανικής μάθησης, όπως τα νευρωνικά δίκτυα και τα δέντρα αποφάσεων, εκπαιδεύονται σε μεγάλα σύνολα δεδομένων για την ταξινόμηση και τον εντοπισμό της σοβαρότητας της ασθένειας.

Η τεχνολογία αυτή επιτρέπει την έγκαιρη διάγνωση, μειώνοντας την εξάρτηση από χημικά φυτοπροστατευτικά και εστιάζοντας στην προληπτική αντιμετώπιση. Παράλληλα, η ενσωμάτωση φασματικών δεδομένων και τεχνητής νοημοσύνης προάγει τη βιώσιμη γεωργία, παρέχοντας στους παραγωγούς εργαλεία ακριβείας για τη διαχείριση των καλλιεργειών τους. Η εφαρμογή αυτών των μεθόδων είναι ιδιαίτερα υποσχόμενη για τη μείωση των περιβαλλοντικών επιπτώσεων και την αύξηση της παραγωγικότητας.

Όργανο Μέτρησης

Η Specim IQ (Εικόνα 1.5.1-2) είναι μια υπερφασματική κάμερα τύπου **push-broom**, σχεδιασμένη για την καταγραφή δεδομένων υψηλής φασματικής και χωρικής ανάλυσης, ιδανική για εφαρμογές ανίχνευσης ασθενειών σε φυτά σε προσυμπτωματικό στάδιο, τόσο στο εργαστήριο όσο και στον αγρό.



Εικόνα 1.5.1-2. Η υπερφασματική κάμερα Specim IQ.

Οι λήψεις και τα χαρακτηριστικά της παρουσιάζονται αναλυτικά παρακάτω:

1. Τύπος Λήψης: Push-Broom Σάρωση

- **Γραμμική καταγραφή (line scan):** Κάθε λήψη της κάμερας συλλέγει μια στενή «λωρίδα» του αντικειμένου. Στη συνέχεια, για να καλυφθεί ολόκληρη η επιφάνεια ενδιαφέροντος, απαιτείται είτε η μετακίνηση της κάμερας είτε του αντικειμένου.
- **Δημιουργία Υπερφασματικού Κύβου (Hyperspectral Cube):** Ο τελικός όγκος δεδομένων περιλαμβάνει:
 - **Χωρικές διαστάσεις (X, Y):** Καταγράφουν το πλάτος και το μήκος της περιοχής ενδιαφέροντος.
 - **Φασματική διάσταση (λ):** Περιλαμβάνει 204 φασματικά κανάλια στο

φάσμα 400–1000 nm (ορατό και εγγύς υπέρυθρο – Vis–NIR).

Αυτό επιτρέπει την καταγραφή του πλήρους φάσματος για κάθε εικονοστοιχείο (pixel), παρέχοντας πολύτιμες πληροφορίες για τις φυσιολογικές ή παθολογικές καταστάσεις των φυτών.

2. Χαρακτηριστικά και Πλεονεκτήματα

Φορητότητα και Ευχρηστία:

- Η Specim IQ είναι μια πλήρως αυτόνομη συσκευή, με ενσωματωμένη μπαταρία, οθόνη αφής και δυνατότητα αποθήκευσης δεδομένων στη δική της μνήμη.
- Η εργονομία της και το βάρος της (1.3 kg) την καθιστούν εύχρηστη για επιτόπιες εφαρμογές στον αγρό (in situ), σε αντίθεση με άλλες υπερφασματικές κάμερες που απαιτούν επιπλέον εξοπλισμό.

Αυτόματες Διορθώσεις:

- Η κάμερα διαθέτει **μηχανισμό αυτόματης λήψης Σκοτεινής Αναφοράς (Dark Reference)**, κλείνοντας το διάφραγμα (shutter) ώστε να εξαλείψει το φασματικό θόρυβο.
- Παρέχει επίσης τη δυνατότητα λήψης **Λευκής Αναφοράς (White Reference)** μέσω της επιλογής "Simultaneous White Reference," που απαιτεί την παρουσία της Λευκής Πλάκας (White Panel) στο οπτικό πεδίο για κάθε λήψη.

Ευέλικτη Καταγραφή και Επεξεργασία Δεδομένων:

- Τα δεδομένα αποθηκεύονται σε **"Default Recording Mode (DRM)"**, διατηρώντας τα πλήρως ακατέργαστα (raw) για μεταγενέστερη ανάλυση.
- Παρέχει επιλογές σύνδεσης μέσω USB ή Wi-Fi, διευκολύνοντας τη μεταφορά δεδομένων σε προσωπικό υπολογιστή για περαιτέρω επεξεργασία.

Ενσωματωμένο GPS:

- Διαθέτει σύστημα ενσωματωμένου GPS, που καταγράφει αυτόματα τις γεωγραφικές συντεταγμένες κάθε λήψης. Αυτό είναι εξαιρετικά χρήσιμο για τη χαρτογράφηση της κατανομής ασθενειών στον αγρό και τη συσχέτιση των δεδομένων με τοπογραφικές συνθήκες.

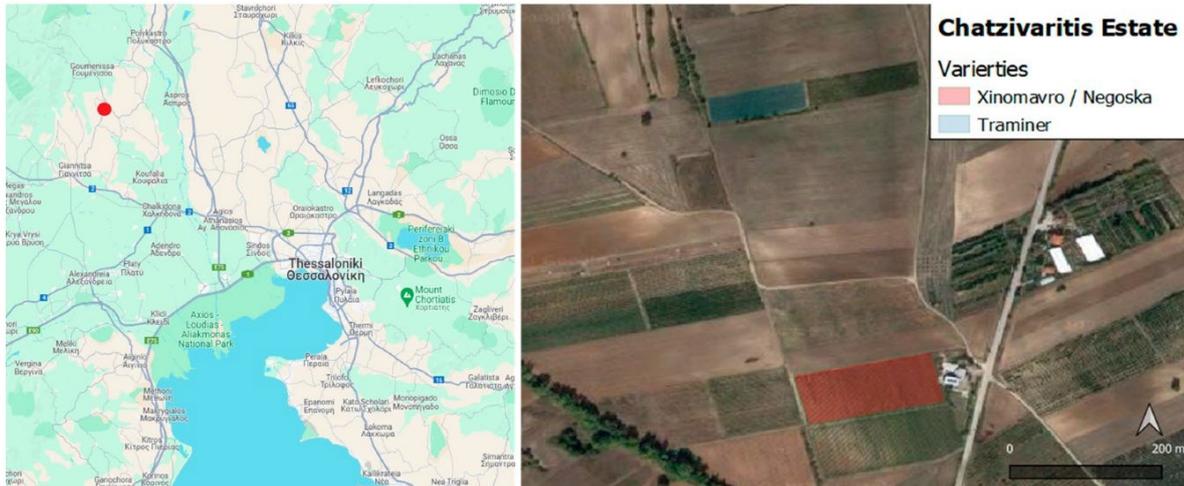
Ευκολία Χρήσης:

- Η ενσωματωμένη οθόνη αφής και το φιλικό λογισμικό της κάμερας επιτρέπουν στους χρήστες να διαχειρίζονται εύκολα τις ρυθμίσεις, να ελέγχουν την ποιότητα των δεδομένων σε πραγματικό χρόνο και να προσαρμόζουν τη διαδικασία λήψης στις ανάγκες τους.

Πειραματική Διαδικασία για την καλλιέργεια αμπελιού Περιοχή μελέτης

Το οινοποιείο Χατζηβαρύτη (1.5.1-3) είναι ένας βιολογικός αμπελώνας που βρίσκεται στην περιοχή της Γουμένισσας, στην περιφέρεια Κεντρικής Μακεδονίας. Το κτήμα εκτείνεται σε 18 εκτάρια γης και βρίσκεται σε γεωγραφικό πλάτος 40°53'21,67"Β και γεωγραφικό μήκος 22°28'37,66"Α. Ο αμπελώνας είναι αφιερωμένος στην καλλιέργεια

και διατήρηση των παραδοσιακών ελληνικών ποικιλιών σταφυλιού όπως το Ξινόμαυρο, η Νεγκόσκα, ο Ροδίτης, η Μαλαγουζιά, το Μοσχάτο και το Ασύρτικο, καθώς και άλλες διεθνείς ποικιλίες όπως Traminer, Sauvignon blanc, Merlot, Chardonnay και Cabernet Sauvignon.



Εικόνα 1.5.1-3. Το οινοποιείο Χατζηβαρίτη, που βρίσκεται στην περιοχή της κ. Μακεδονίας (αριστερά), με τα πεδία έρευνας (δεξιά). Το χωράφι των ποικιλιών Ξινόμαυρο και Νεγκόσκα βρίσκεται στα νότια, ενώ το χωράφι με την ποικιλία Traminer βρίσκεται βόρεια του οινοποιείου.

Το κτήμα Χατζηβαρύτη είναι ιδιαίτερα γνωστό για τη δέσμευσή του σε πρακτικές βιώσιμης γεωργίας και στη χρήση μεθόδων βιολογικής γεωργίας.

Οι λήψεις έγιναν σε τμήματα του αμπελώνα Χατζηβαρύτη σε αμπέλια των ποικιλιών Ξινόμαυρο και Νεγκόσκα (Εικόνα 1.5.1-4) και Traminer (Εικόνα 1.5.1-5). Οι επιλεγμένες φωτογραφίες ελήφθησαν τόσο από υγιή όσο και από φυτά μολυσμένα από τη μυκητολογική ασθένεια του ωιδίου.



Εικόνα 1.5.1-4. Τμήμα του κτήματος Χατζηβαρύτη, που απεικονίζει τον αγρό με τις ποικιλίες Ξινόμαυρο και Νεγκόσκα.



Εικόνα 1.5.1-5. Τμήμα του κτήματος Χατζηβαρύτη, που απεικονίζει τον αγρό με την ποικιλία Traminer

Ταυτοποίηση παθογόνου και σήμανση δειγμάτων

Η αξιολόγηση του αμπελώνα πραγματοποιήθηκε σε πραγματικές συνθήκες αγρού, χωρίς τεχνητή μόλυνση. Έμπειρος φυτοπαθολόγος του κτήματος, με εξειδίκευση στην αμπελοκαλλιέργεια, εντόπισε φύλλα με ορατά συμπτώματα ωιδίου, τα οποία και επισημάνθηκαν.

Λαμβάνοντας υπόψη ότι ο μύκητας (*Erysiphe necator*) παράγει σπόρια (κονίδια) που μπορούν να μεταφερθούν μέσω ανέμου ή ακόμα και μέσω άμεσης επαφής μεταξύ φύλλων, θεωρήθηκε πολύ πιθανό ότι τα γειτονικά φύλλα των ήδη μολυσμένων θα μολυνθούν μελλοντικά.

Για τον λόγο αυτό, επιλέχθηκαν διπλανά φύλλα που αρχικά θεωρούνταν υγιή (χωρίς ορατά συμπτώματα), με την εκτίμηση ότι η ασθένεια θα μπορούσε να μεταφερθεί σε αυτά σε σύντομο χρονικό διάστημα, λόγω της χωρικής εγγύτητας με τα μολυσμένα φύλλα.

Ταυτόχρονα, έγινε συλλογή δεδομένων από πραγματικά υγιή φύλλα (σε περιοχές του αμπελώνα όπου δεν υπήρχαν ενδείξεις μόλυνσης και όπου δεν είχε καταγραφεί ιστορικό εμφάνισης ωιδίου).

Διαδικασία λήψης

Τα υπερφασματικά δεδομένα συλλέχθηκαν με τη χρήση της κάμερας Specim IQ. Οι αισθητήρες υπερφασματικής απεικόνισης (HSI) κατασκευάζονται συνήθως για χρήση σε ελεγχόμενα, εργαστηριακά περιβάλλοντα. Για τον λόγο αυτό, οι κατασκευαστές παρέχουν συνήθως επιλογές για την απόκτηση κατάλληλων λήψεων. Μια επιτυχημένη έρευνα πεδίου αυτού του τύπου πρέπει να λαμβάνει υπόψη τις συνθήκες φωτισμού. Ως εκ τούτου, οι έρευνες πραγματοποιούνται κατά τους καλοκαιρινούς μήνες, υπό ηλιόλουστες συνθήκες.

Οι χειριστές της κάμερας πρέπει να τοποθετούν τον εξοπλισμό με τρόπο ώστε να αποφεύγεται το άμεσο ηλιακό φως, το οποίο μπορεί να επηρεάσει αρνητικά την ποιότητα της λήψης και να προκαλέσει βλάβη στον αισθητήρα. Το πρώτο βήμα είναι η σωστή βαθμονόμηση του αισθητήρα (Εικόνα 1.5.1-6). Ο εξοπλισμός της Specim IQ περιλαμβάνει ένα πορτοκαλί "πλακίδιο βαθμονόμησης" (calibration tile) και μία πλάκα λευκής αναφοράς (white reference panel), που χρησιμοποιούνται για τη βαθμονόμηση. Οι χειριστές πρέπει να αξιολογούν την ποιότητα της λήψης βάσει των τιμών ανακλαστικότητας που εμφανίζονται στην οθόνη της συσκευής και στη συνέχεια να προχωρούν στη δημιουργία του υπερφασματικού συνόλου δεδομένων (HSI dataset).



Εικόνα 1.5.1-6. Διαδικασία βαθμονόμησης της υπερφασματικής κάμερας υπό συνθήκες ανοιχτού αγρού.

Η συσκευή τοποθετείται σε απόσταση περίπου 60 cm από το φυτό-στόχο. Κάθε λήψη μπορεί να διαρκέσει από λίγα δευτερόλεπτα έως λίγα λεπτά, ανάλογα με τις συνθήκες φωτισμού, χωρίς ωστόσο να επηρεάζεται η ποιότητα της εικόνας. Η πλάκα λευκής αναφοράς πρέπει να συμπεριλαμβάνεται στη "λειτουργία ταυτόχρονης αναφοράς" (simultaneous mode), η οποία παρέχει επιλογές για λήψη της λευκής αναφοράς μία φορά πριν από κάθε μέτρηση ή για χρήση προκαθορισμένης λευκής αναφοράς από τη συσκευή. Παρόλα αυτά, η λειτουργία ταυτόχρονης αναφοράς πρέπει να χρησιμοποιείται όπου είναι δυνατόν για καλύτερη ποιότητα δεδομένων. Κάθε λήψη αξιολογείται αρχικά επί τόπου μέσω της οθόνης της συσκευής μετά από κάθε καταγραφή.

Οι έρευνες πραγματοποιήθηκαν σε συγκεκριμένα φυτά βάσει σχολίων από τους εργατές του οινοποιείου. Ένα μεγάλο σύνολο 112 υπερφασματικών εικόνων καταγράφηκε, εκ των οποίων οι 81 προέρχονται από φυτά μολυσμένα με ωίδιο. Τα δεδομένα εξήχθησαν σε βιβλιοθήκες ανάλυσης μέσω του λογισμικού Specim IQ Studio, όπου αξιολογήθηκαν περαιτέρω για την ποιότητά τους.

Πειραματική Διαδικασία για την καλλιέργεια τομάτας

Πειραματική Διάταξη

Το πείραμα επικεντρώθηκε στη μελέτη της αντίδρασης της ποικιλίας τομάτας *Ελπίδα* σε τεχνητές μολύνσεις με τον μύκητα *Botrytis cinerea*, χρησιμοποιώντας προηγμένες μεθόδους ανίχνευσης, όπως η υπερφασματική απεικόνιση. Για την εκτέλεση του πειράματος χρησιμοποιήθηκαν φυτάρια ηλικίας τριών εβδομάδων, τα οποία προμηθεύτηκαν από την εταιρεία Agris. Τα φυτάρια μεταφυτεύθηκαν σε πλαστικά γλαστράκια διαστάσεων 30 × 30 × 30 cm, με εδαφικό υπόστρωμα που περιείχε 30% περλίτη και 70% φυτόχωμα. Μετά τη μεταφύτευση, τα φυτά αναπτύχθηκαν σε ελεγχόμενο θάλαμο ανάπτυξης υπό συνθήκες θερμοκρασίας 22°C, σχετικής υγρασίας >85% και φωτοπεριόδου 16:8 (φως/σκοτάδι) (Εικόνα 1.5.1-7). Συνολικά καλλιεργήθηκαν 120 φυτά, εκ των οποίων το 30% χρησιμοποιήθηκε ως μάρτυρες (μη μολυσμένα φυτά), ενώ το υπόλοιπο 70% υποβλήθηκε σε τεχνητή μόλυνση.



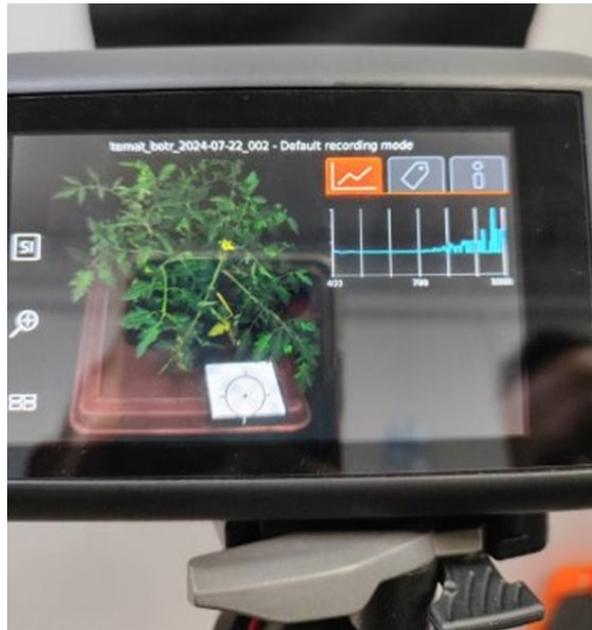
Εικόνα 1.5.1-7. Θάλαμος ελεγχόμενων συνθηκών ανάπτυξης των νεαρών φυτών τομάτας του εργαστηρίου Φυτοπαθολογίας του τμήματος Γεωπονίας του Α.Π.Θ.

Η προετοιμασία του μολύσματος πραγματοποιήθηκε με το στέλεχος αναφοράς *Botrytis cinerea* (B.05). Το στέλεχος καλλιεργήθηκε σε θρεπτικό υπόστρωμα Potato Dextrose Agar (PDA) σε τρυβλία Petri διαμέτρου 90 mm, τα οποία επώασησαν σε θάλαμο στους 20°C υπό σκοτάδι για 15 ημέρες. Μετά την επώαση, τα κονίδια συλλέχθηκαν με αποστειρωμένο βαμβακοφόρο στυλεό και επαναδιαλύθηκαν σε διάλυμα Gamborg B5. Η συγκέντρωση του αιωρήματος προσδιορίστηκε με τη χρήση αιματοκυτταρόμετρου Neubauer στα 10^5 κονίδια/ml. Τα φυτά μολύνθηκαν τεχνητά μέσω ψεκασμού με το αιώρημα κονιδίων μέχρι απορροής. Αμέσως μετά, τα μολυσμένα φυτά μεταφέρθηκαν στον θάλαμο ανάπτυξης, όπου διατηρήθηκαν σε θερμοκρασία 21°C, σχετική υγρασία >90% και φωτοπερίοδο 16:8.

Διαδικασία Λήψης Υπερφασματικών απεικονίσεων

Η καταγραφή της λοίμωξης πραγματοποιήθηκε με τη χρήση της ίδιας υπερφασματικής κάμερας που χρησιμοποιήθηκε και στην πειραματική διαδικασία των αμπελιών (Specim IQ). Η εστιακή απόσταση της κάμερας ορίστηκε στα 70 cm, με γωνία λήψης 145°, ενώ

τα πάνω μέρη των φύλλων τοποθετήθηκαν στραμμένα προς την κάμερα, δίνοντας έμφαση στο σημείο μόλυνσης (Εικόνα 1.5.1-9). Παράλληλα έγιναν οι κατάλληλες ρυθμίσεις στην κάμερα, μέσω της οθόνης αφής (Εικόνα 1.5.1-8). Το πείραμα αποσκοπούσε στη διερεύνηση των φυσιολογικών αλλαγών που προκύπτουν από τη μόλυνση, στη συγκριτική αξιολόγηση μολυσμένων και μη μολυσμένων φυτών, καθώς και στην αξιολόγηση της αποτελεσματικότητας της υπερφασματικής τεχνολογίας για την έγκαιρη ανίχνευση της λοίμωξης.



Εικόνα 1.5.1-8: Απεικόνιση μέσα από την οθόνη αφής της υπερφασματικής κάμερας.



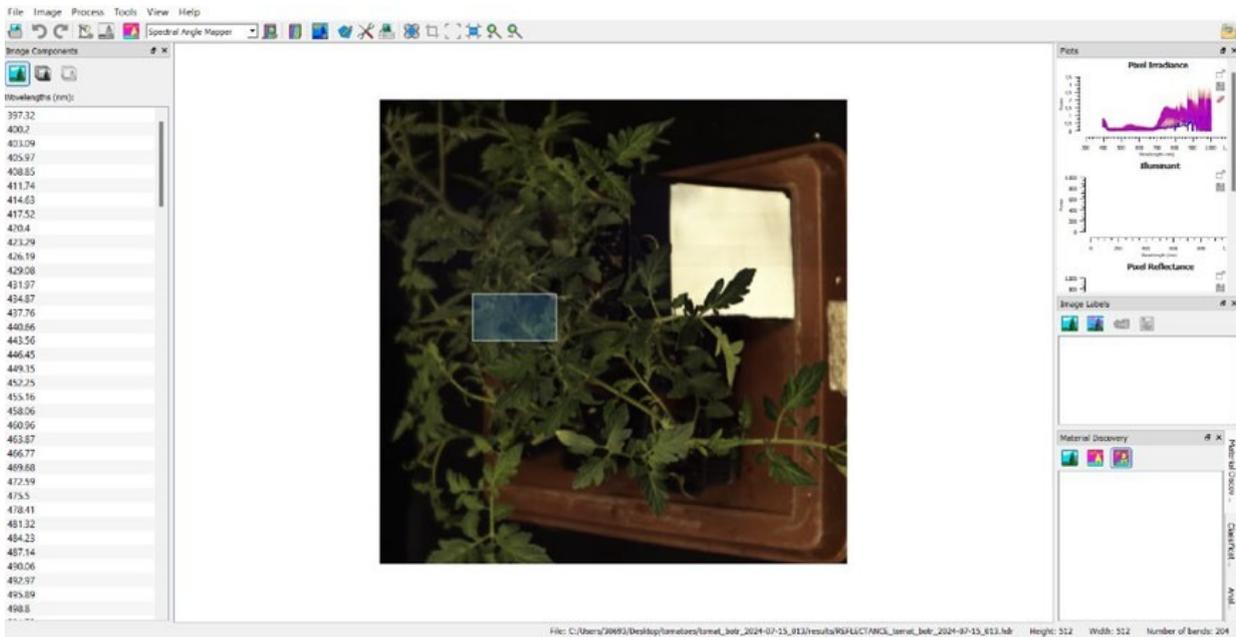
Εικόνα 1.5.1-9. Πειραματική διάταξη για τη λήψη υπερφασματικών απεικονίσεων, με την υπερφασματική κάμερα Specim IQ.

2.2 Αποτελέσματα και Συζήτηση

Εξαγωγή Δεδομένων

Η διαδικασία ξεκίνησε με τη λήψη υπερφασματικών δεδομένων χρησιμοποιώντας την κάμερα της Specim. Πριν τη λήψη, πραγματοποιήθηκε ρύθμιση του εξοπλισμού για την επίτευξη της βέλτιστης λειτουργίας. Η κάμερα προσαρμόστηκε στις απαιτήσεις του εκάστοτε πειράματος, λαμβάνοντας υπόψη την ανάλυση, το εύρος των μηκών κύματος και τις συνθήκες φωτισμού. Τα υπερφασματικά δεδομένα καταγράφηκαν σε μορφές αρχείων όπως ENVI (.hdr) ή TIFF (.tif).

Στη συνέχεια, τα δεδομένα εισήχθησαν στο λογισμικό Skyven για επεξεργασία και ανάλυση (Εικόνα 1.5.1-10). Στο στάδιο της προεπεξεργασίας, εφαρμόστηκαν μέθοδοι διόρθωσης θορύβου για τη βελτίωση της ποιότητας των δεδομένων, καθώς και γεωμετρικές διορθώσεις για την ορθή ευθυγράμμιση των υπερφασματικών εικόνων. Ακολούθως, πραγματοποιήθηκε εξαγωγή των φασματικών υπογραφών από συγκεκριμένες περιοχές ενδιαφέροντος (Regions of Interest - ROI). Η διαδικασία αυτή περιλάμβανε την επιλογή των περιοχών, την ανάλυση των φασματικών χαρακτηριστικών και την εξαγωγή δεδομένων, όπως η ανακλαστικότητα και τα χαρακτηριστικά απορρόφησης.



Εικόνα 1.5.1-10. Απεικόνιση του περιβάλλοντος του λογισμικού Skysen, για την ανάλυση των υπερφασματικών εικόνων και την επιλογή των σημείων ενδιαφέροντος.

Μετά την επεξεργασία, τα αποτελέσματα εξήχθησαν σε μορφή CSV. Για την εξαγωγή αυτή, χρησιμοποιήθηκε η λειτουργία εξαγωγής του Skysen, όπου επιλέχθηκαν τα δεδομένα που έπρεπε να αποθηκευτούν, όπως τα μήκη κύματος και οι αντίστοιχες τιμές ανακλαστικότητας. Ορίστηκε η δομή του αρχείου CSV, περιλαμβάνοντας κεφαλίδες και διαχωριστικά, και τα δεδομένα αποθηκεύτηκαν σε μορφή κατάλληλη για ανάλυση με άλλα εργαλεία. Τέλος, πραγματοποιήθηκε επαλήθευση της ακρίβειας των αποθηκευμένων δεδομένων, διασφαλίζοντας την ποιότητα και τη χρησιμότητά τους.

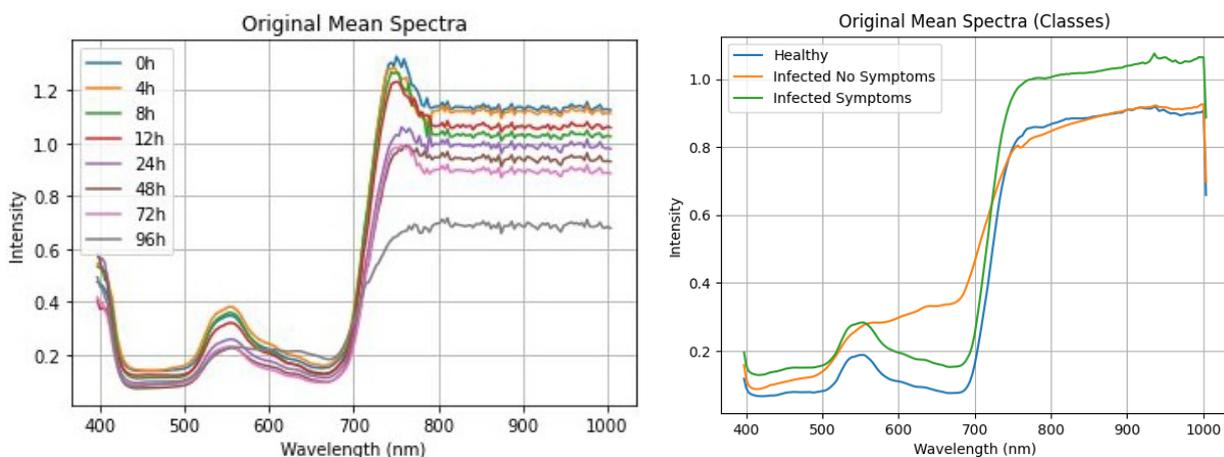
A	B	C	D	E	F	G	H
image	x	y	label	397.32	400.2	403.09	405.97
REFLECTA	30	393	8	0.25	0.25	0.333333	0.25
REFLECTA	31	393	8	0.333333	0.25	0.222222	0.25
REFLECTA	28	394	8	0.222222	0.333333	0.333333	0.333333
REFLECTA	29	394	8	0.222222	0.222222	0.333333	0.333333
REFLECTA	30	394	8	0.125	0.25	0.333333	0.25
REFLECTA	31	394	8	0.222222	0.125	0.222222	0.25
REFLECTA	27	395	8	0.25	0.333333	0.125	0.222222
REFLECTA	28	395	8	0.222222	0.222222	0.222222	0.333333
REFLECTA	29	395	8	0.222222	0.333333	0.333333	0.333333
REFLECTA	30	395	8	0.25	0.25	0.222222	0.25
REFLECTA	31	395	8	0.333333	0.25	0.222222	0.25
REFLECTA	25	396	8	0.333333	0.25	0.333333	0.333333
REFLECTA	26	396	8	0.333333	0.333333	0.125	0.333333
REFLECTA	27	396	8	0.25	0.222222	0.25	0.333333
REFLECTA	28	396	8	0.333333	0.333333	0.222222	0.222222
REFLECTA	29	396	8	0.333333	0.333333	0.333333	0.333333

Εικόνα 1.5.1-11. Στιγμιότυπο του αρχείου .csv που δημιουργείται μετά την εξαγωγή

των φασματικών δεδομένων από το λογισμικό Skyven, με τις τιμές αντανάκλασης και κάποια μεταδεδομένα.

Τα δεδομένα που συλλέχθηκαν περιλάμβαναν φασματικές υπογραφές από τις επιλεγμένες περιοχές, οι οποίες αναλύθηκαν περαιτέρω για την εξαγωγή χρήσιμων πληροφοριών. Μέσω της ανάλυσης αυτής, εντοπίστηκαν χαρακτηριστικά απορρόφησης και ανακλαστικότητας, που είναι κρίσιμα για την κατανόηση της συμπεριφοράς των δειγμάτων.

Τα δεδομένα, στη συνέχεια χωρίστηκαν στις κλάσεις ανάλογα με την πειραματική διαδικασία που ακολουθήθηκε για το κάθε φυτό. Έτσι, για το πείραμα με την τομάτα, δημιουργήθηκαν 8 κλάσεις, βάσει της ώρας κατά την οποία έγινε η δειγματοληψία. Έτσι, σχηματίστηκαν κλάσεις για την έναρξη του πειράματος (0h) και για την κάθε φάση λήψης (4h, 8h, 12h, 24h, 48h, 72h, 96h). Αντίστοιχα, στην περίπτωση του πειράματος που έγινε για την καλλιέργεια αμπελιού, σχηματίστηκαν τρεις κλάσεις, οι οποίες αντιπροσώπευαν την υγιή κατάσταση, τη μολυσμένη, χωρίς ορατά συμπτώματα της ασθένειας και τη μολυσμένη, με ορατά τα συμπτώματα από την ασθένεια. Τα διαγράμματα από τους μέσους όρους της κάθε κλάσης τόσο για το πείραμα στην καλλιέργεια τομάτας και την καλλιέργεια σε αμπέλια πειράματος φαίνονται στην Εικόνα 1.5.1-12, αριστερά και δεξιά αντίστοιχα.



Εικόνα 1.5.1-12. Μέσος όρος των επιλεγμένων φασματικών αποκρίσεων για κάθε μία από τις 8 κλάσεις των φυτών τομάτας που αφορούν στο χρόνο που πέρασε μετά τη μόλυνση των φυτών (αριστερά) και για τις 3 κλάσεις των φυτών αμπελού που αφορούν στη φαινοτυπική κατάσταση υγείας των φυτών.

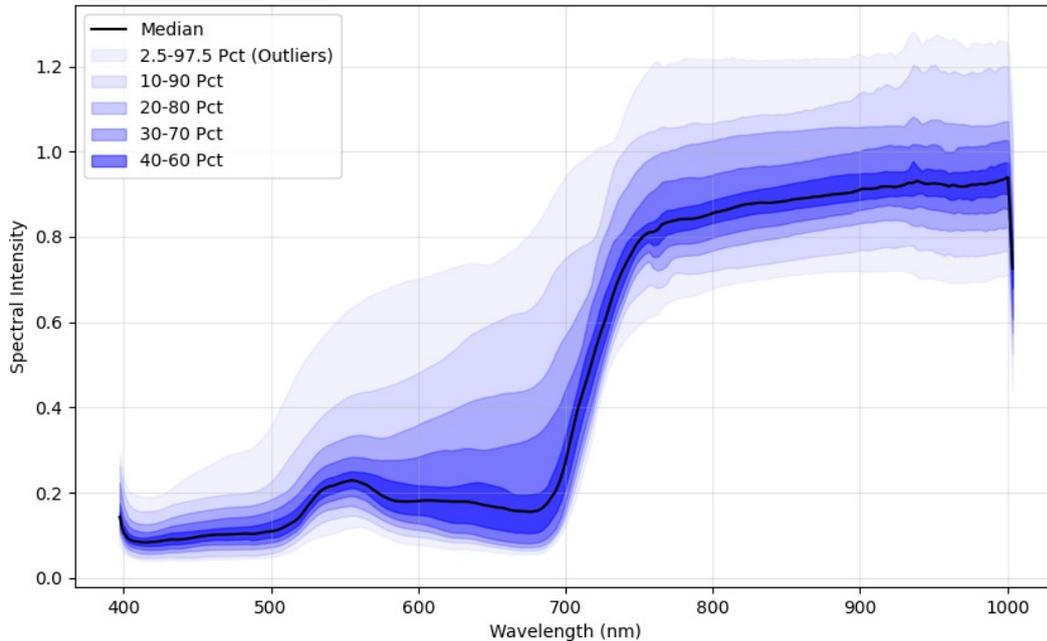
Προεπεξεργασία Φασματικών Δεδομένων

Ακολουθήθηκε η ίδια διαδικασία προεπεξεργασίας δεδομένων και στα δύο πειράματα (αμπέλι και τομάτα), εφόσον αυτή αφορούσε την επεξεργασία υπερφασματικών εικόνων.

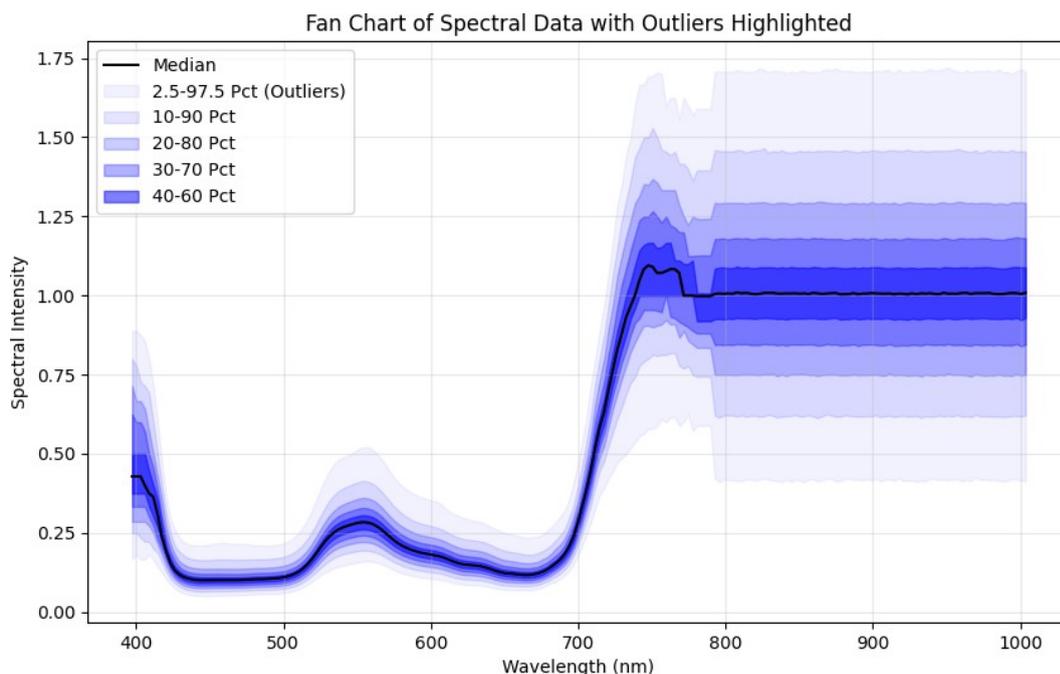
Η προεπεξεργασία φασματικών δεδομένων αποτελεί κρίσιμο βήμα για τη μείωση του θορύβου που προκύπτει από το όργανο μέτρησης και τη διευκόλυνση της επεξεργασίας τους από αλγορίθμους σε επόμενο στάδιο.

Σε πρώτη φάση, γίνεται αφαίρεση των ακραίων τιμών, που αφορά στο 2.5% των

δεδομένων από τις κορυφαιές και χαμηλότερες φασματικές υπογραφές. Οι τιμές που έχουν απαλειφθεί φαίνονται στις παρακάτω εικόνες για το αμπέλι (Εικόνα 1.5.1-13) και την τομάτα (Εικόνα 1.5.1-14).



Εικόνα 1.5.1-13. Διάγραμμα που απεικονίζει τα διαφορετικά εκατοστημόρια του συνόλου των φασματικών δεδομένων από τα φυτά αμπέλου που χρησιμοποιήθηκαν στην παρούσα μελέτη, καθώς και το διάμεσο με μαύρη γραμμή.

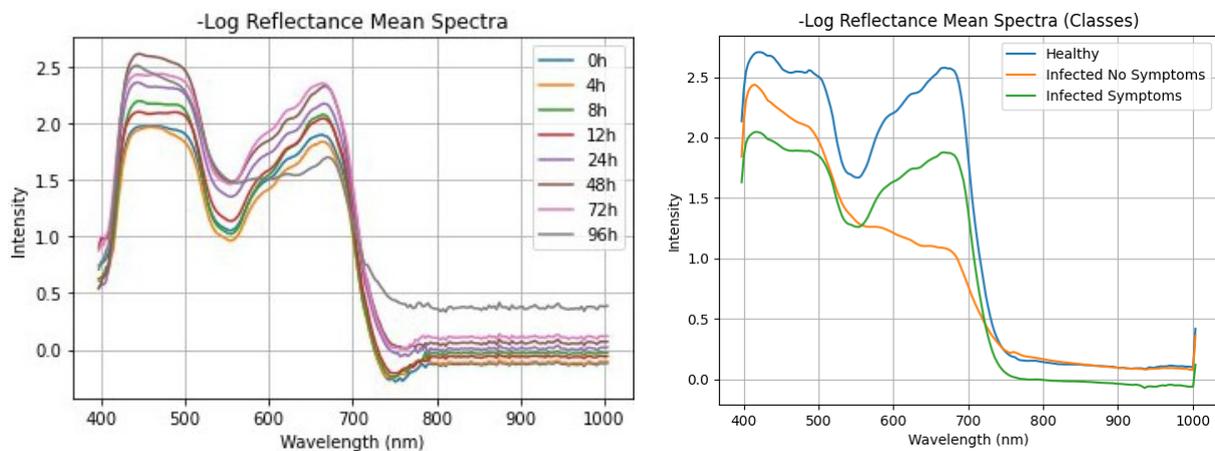


Εικόνα 1.5.1-14. Διάγραμμα που απεικονίζει τα διαφορετικά εκατοστημόρια του συνόλου των φασματικών δεδομένων από τα φυτά τομάτας που χρησιμοποιήθηκαν στην παρούσα μελέτη,

καθώς και το διάμεσο με μαύρη γραμμή.

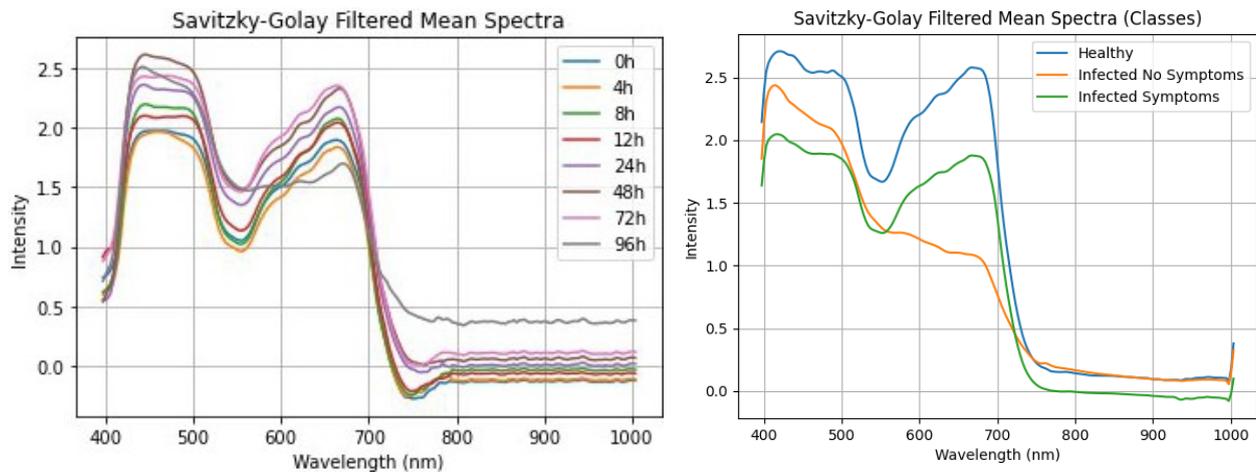
Έτσι, μετά την αφαίρεση των ακραίων τιμών, δημιουργήθηκε η δομοσειρά (pipeline) προεπεξεργασίας των δεδομένων με τη σειρά που παρατίθεται παρακάτω:

1. Αρνητικός λογάριθμος Η χρήση του αρνητικού λογαρίθμου βοηθά στη μετατροπή των φασματικών δεδομένων σε λογαριθμική κλίμακα, μειώνοντας τις μη γραμμικότητες και ενισχύοντας τα χαρακτηριστικά του φάσματος που είναι λιγότερο εμφανή στη γραμμική κλίμακα. Αυτή η μέθοδος είναι ιδιαίτερα χρήσιμη σε φάσματα απορρόφησης, όπως αυτά που προκύπτουν από την υπέρυθη φασματοσκοπία (IR), καθώς συμβάλλει στην αναγωγή του θορύβου από την απόκριση του οργάνου μέτρησης (Rinnan et al., 2009). Αποτελέσματα του αρνητικού αλγορίθμου στα δεδομένα των δύο πειραμάτων φαίνονται στην Εικόνα 1.5.1-15.



Εικόνα 1.5.1-15. Προεπεξεργασία των μέσων όρων των φασματικών αποκρίσεων με τον αρνητικό λογάριθμο, για κάθε μία από τις 8 κλάσεις των φυτών της τομάτας, που αφορούν στο χρόνο από τη μόλυνση των φυτών (αριστερά) και για τις 3 κλάσεις των φυτών αμπέλου, που αφορούν στη φαινοτυπική κατάσταση της υγείας των φυτών (δεξιά).

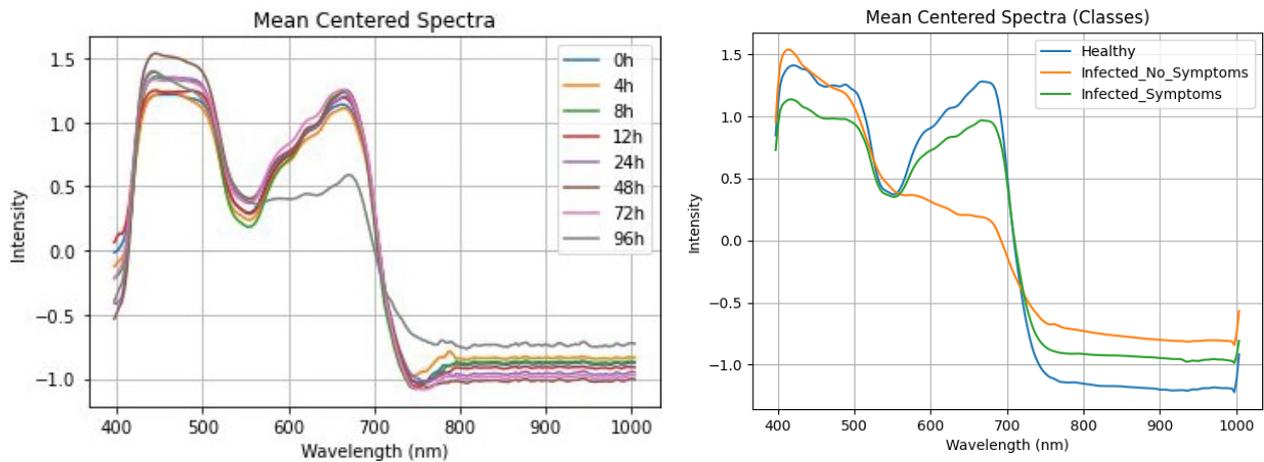
2. Φίλτρο Savitzky-Golay Το φίλτρο Savitzky-Golay διαδέχεται τον αρνητικό αλγόριθμο σε σειρά στην προεπεξεργασία των δεδομένων αυτής της έρευνας και αποτελεί μία τεχνική εξομάλυνσης, η οποία διατηρεί την ακρίβεια των κορυφών και άλλων σημαντικών χαρακτηριστικών του φάσματος. Στη συγκεκριμένη περίπτωση, χρησιμοποιήθηκε πολυώνυμο δευτέρου βαθμού με 5 σημεία υποστήριξης (window size), το οποίο εξισορροπεί την απομάκρυνση του θορύβου και τη διατήρηση της γεωμετρίας των δεδομένων. Αυτή η μέθοδος είναι ιδανική για την εξάλειψη του τυχαίου θορύβου χωρίς να αλλοιώνεται η ποιότητα του σήματος (Savitzky & Golay, 1964). Αποτελέσματα του φίλτρου στα δεδομένα των δύο πειραμάτων φαίνονται στην Εικόνα 1.5.1-16.



Εικόνα 1.5.1-16. Προεπεξεργασία των μέσων όρων των προεπεξεργασμένων με τον αρνητικό αλγόριθμο φασματικών αποκρίσεων με το φίλτρο Savitzky-Golay (2^{ου} βαθμού πολυώνυμο, με 5 σημεία υποστήριξης), για κάθε μία από τις 8 κλάσεις των φυτών της τομάτας, που αφορούν στο χρόνο από τη μόλυνση των φυτών (αριστερά) και για τις 3 κλάσεις των φυτών αμπέλου, που αφορούν στη φαινοτυπική κατάσταση της υγείας των φυτών (δεξιά).

3. Mean Centering (Κεντροποίηση ως προς το μέσο όρο) Η κεντροποίηση ως προς το μέσο όρο αποτελεί την τελευταία μέθοδο προεπεξεργασίας των πειραματικών δεδομένων αυτής της έρευνας και είναι μία μέθοδος που κανονικοποιεί τα δεδομένα αφαιρώντας τον μέσο όρο από κάθε φασματική τιμή. Με αυτόν τον τρόπο, μειώνεται η επίδραση της βασικής γραμμής (baseline) και διευκολύνεται η σύγκριση μεταξύ διαφορετικών δειγμάτων. Η μέθοδος αυτή χρησιμοποιείται συχνά πριν από την εφαρμογή τεχνικών ανάλυσης, όπως η PCA (Principal Component Analysis), για την εξαγωγή κρυφών πληροφοριών από τα φάσματα (Martens & Næs, 1989). Αποτελέσματα του mean centering στα δεδομένα των δύο πειραμάτων φαίνονται στην Εικόνα 1.5.1.-17. Από μία σύγκριση αυτής της εικόνας με τις αντίστοιχες εικόνες από τις προηγούμενες μεθόδους προεπεξεργασίας, φαίνεται ότι οι επιλεγμένες φασματικές υπογραφές έχουν κανονικοποιηθεί και εξομαλυνθεί σημαντικά. Αυτό γίνεται ιδιαίτερα προφανές στην περίπτωση των δεδομένων της τομάτας.

Η συνδυαστική χρήση αυτών των μεθόδων προεπεξεργασίας μπορεί να οδηγήσει σε δεδομένα υψηλής ποιότητας, βελτιώνοντας σημαντικά την ακρίβεια και την αποτελεσματικότητα των επόμενων σταδίων ανάλυσης.



Εικόνα 1.5.1-17. Προεπεξεργασία των μέσων όρων των προεπεξεργασμένων με τη μέθοδο mean centering φασματικών αποκρίσεων που έχουν υποστεί επεξεργασία με τον αρνητικό λογάριθμο και το φίλτρο Savitzky-Golay για κάθε μία από τις 8 κλάσεις των φυτών της τομάτας, που αφορούν στο χρόνο από τη μόλυνση των φυτών (αριστερά) και των 3 κλάσεων των φυτών αμπέλου, που αφορούν στη φαινοτυπική κατάσταση της υγείας των φυτών (δεξιά).

Τέλος, όλα τα δεδομένα, πριν εισαχθούν στα μοντέλα μηχανικής μάθησης έχουν υποστεί τη διαδικασία της υποδειγματοληψίας (undersampling), ώστε να υπάρχει αντιπροσωπευτικότητα σε κάθε διαφορετική κλάση, τόσο στα αμπέλια όσο και στις τομάτες. Η μέθοδος υποδειγματοληψίας που εφαρμόστηκε είναι η τυχαία υποδειγματοληψία. μέθοδος αυτή είναι ιδανική για τη διατήρηση της ποικιλίας των φασματικών σημάτων, ενώ μειώνει τον όγκο των δεδομένων που απαιτείται για ανάλυση, εξοικονομώντας χρόνο και υπολογιστικούς πόρους. Εξασφαλίζει επίσης την ανεξαρτησία και τη στατιστική εγκυρότητα των δεδομένων, διευκολύνοντας την περαιτέρω ανάλυση και την ανάπτυξη ακριβών μοντέλων.

Εξαγωγή και Επιλογή Χαρακτηριστικών

Εξαγωγή Χαρακτηριστικών μέσω Δεικτών Βλάστησης

Οι δείκτες βλάστησης είναι μετασχηματισμοί δύο ή περισσότερων φασματικών ζωνών, συνδυασμένων έτσι ώστε να βελτιώνεται η συνεισφορά των ιδιοτήτων της βλάστησης στην εξαγωγή συμπερασμάτων από την μελέτη του φάσματος (Huet et al., 2002). Οι δείκτες βλάστησης παρουσιάζουν μεγαλύτερη ευαισθησία σε σχέση με τις μεμονωμένες φασματικές ζώνες (Bannari et al., 1995).

Στην παρούσα περίπτωση επιλέχθηκαν αρχικά 12 διαφορετικοί δείκτες βλάστησης (Πίνακας 1.5.1-1) με βάση τις φασματικές ζώνες που χρησιμοποιούνται στον κάθε δείκτη.

Πίνακας 1.5.1-1. Φασματικοί Δείκτες Βλάστησης που χρησιμοποιήθηκαν την παρούσα έρευνα για την εξαγωγή χαρακτηριστικών από τις φασματικές αποκρίσεις των φυτών αμπέλου και τομάτας.

α/α	Δείκτης Βλάστησης	Τύπος Δείκτη Βλάστησης	Βιβλιογραφία
1	NDVI	$\frac{R_{800} - R_{670}}{R_{800} + R_{670}}$	(Rouse et al., 1972)
2	REHBI	$\frac{(R_{842} - R_{665})(842 - 665) - (R_{783} - R_{665})(783 - 665)}{2}$	(Liu et al., 2020)
3	NDRE	$\frac{R_{790} - R_{720}}{R_{790} + R_{720}}$	(Barnes et al., 2000)
4	TVI	$0,5 * [120 * (R_{750} - R_{550}) - 200 * (R_{670} - R_{550})]$	(Broge & Leblanc, 2001)
5	ARI	$\frac{1}{R_{550}} - \frac{1}{R_{700}}$	(Gitelson et al., 2003)
6	gNDVI	$\frac{R_{810} - R_{560}}{R_{810} + R_{560}}$	(Datt, 1998)
7	PSSRa	$\frac{R_{800}}{R_{675}}$	(Blackburn, 1998)
8	PSSRb	$\frac{R_{800}}{R_{650}}$	(Blackburn, 1998)
9	REIP	$700 + 40 * \frac{R_{RE} - R_{700}}{R_{740} - R_{700}}, \text{ όπου } R_{RE} = \frac{R_{670} + R_{780}}{2}$	(Dawson & Curran, 1998)
10	SR _{CHLTOT}	$\frac{R_{760}}{R_{500}}$	(Datt, 1998)
11	VOG1	$\frac{R_{740}}{R_{720}}$	(Vogelmann et al., 1993)
12	VOG2	$\frac{R_{734} - R_{747}}{R_{715} - R_{720}}$	(Vogelmann et al., 1993)

Επιλογή Χαρακτηριστικών

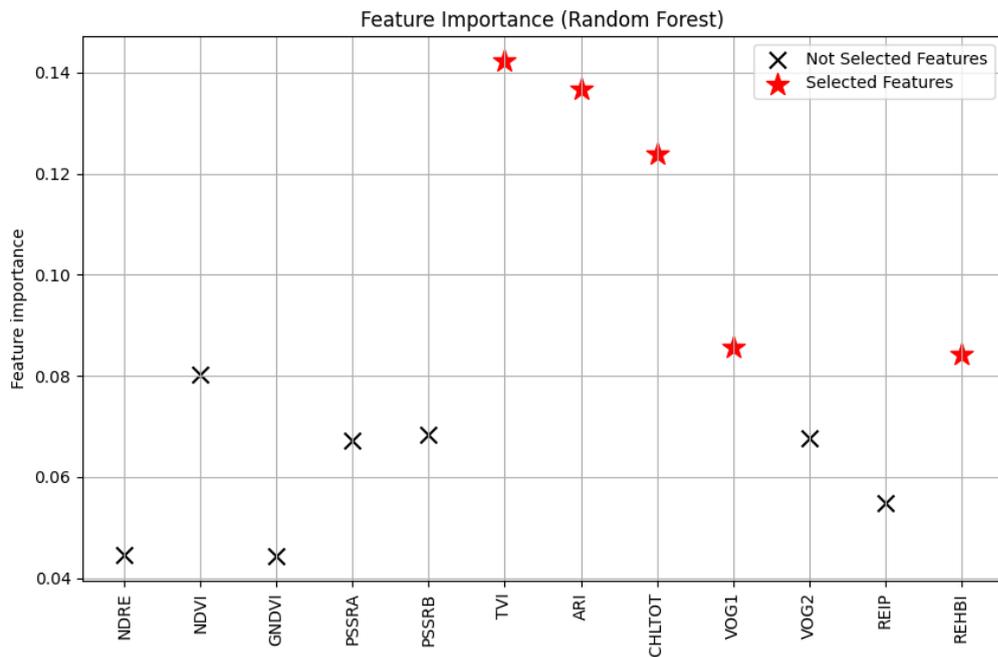
Η επιλογή χαρακτηριστικών αποτελεί ένα κρίσιμο βήμα στη διαδικασία της ανάλυσης δεδομένων και της ανάπτυξης μοντέλων μηχανικής μάθησης. Στην προκειμένη περίπτωση, χρησιμοποιήθηκε το μοντέλο Τυχαία Δάση (Random Forest, RF), το οποίο προσφέρει έναν αξιόπιστο τρόπο για την αξιολόγηση της σημαντικότητας των χαρακτηριστικών. Το μοντέλο Random Forest είναι ένα σύνολο από δέντρα απόφασης που συνεργάζονται για να παράγουν προβλέψεις. Για την αξιολόγηση της σημαντικότητας των χαρακτηριστικών, το μοντέλο εκπαιδεύεται με το πλήρες σύνολο δεδομένων, δοκιμάζοντας πολλούς συνδυασμούς χαρακτηριστικών. Κατά την εκπαίδευση, κάθε δέντρο απόφασης κατασκευάζεται χρησιμοποιώντας τυχαία δείγματα δεδομένων και υποσύνολα χαρακτηριστικών, γεγονός που ενισχύει την ανθεκτικότητα και την ακρίβεια του μοντέλου.

Μετά την εκπαίδευση, η σημαντικότητα κάθε χαρακτηριστικού υπολογίζεται με τη μέθοδο Mean Decrease Impurity (MDI). Η μέθοδος αυτή μετρά τη μέση μείωση της αβεβαιότητας (impurity) που προκαλείται από τη χρήση ενός συγκεκριμένου χαρακτηριστικού για τον διαχωρισμό των δεδομένων στα δέντρα απόφασης. Η αβεβαιότητα μπορεί να μετρηθεί με τη μέθοδο Gini Impurity ή Entropy, ανάλογα με την παράμετρο που έχει οριστεί κατά την εκπαίδευση. Όσο μεγαλύτερη είναι η συμβολή ενός χαρακτηριστικού στη μείωση της αβεβαιότητας, τόσο πιο σημαντικό θεωρείται.

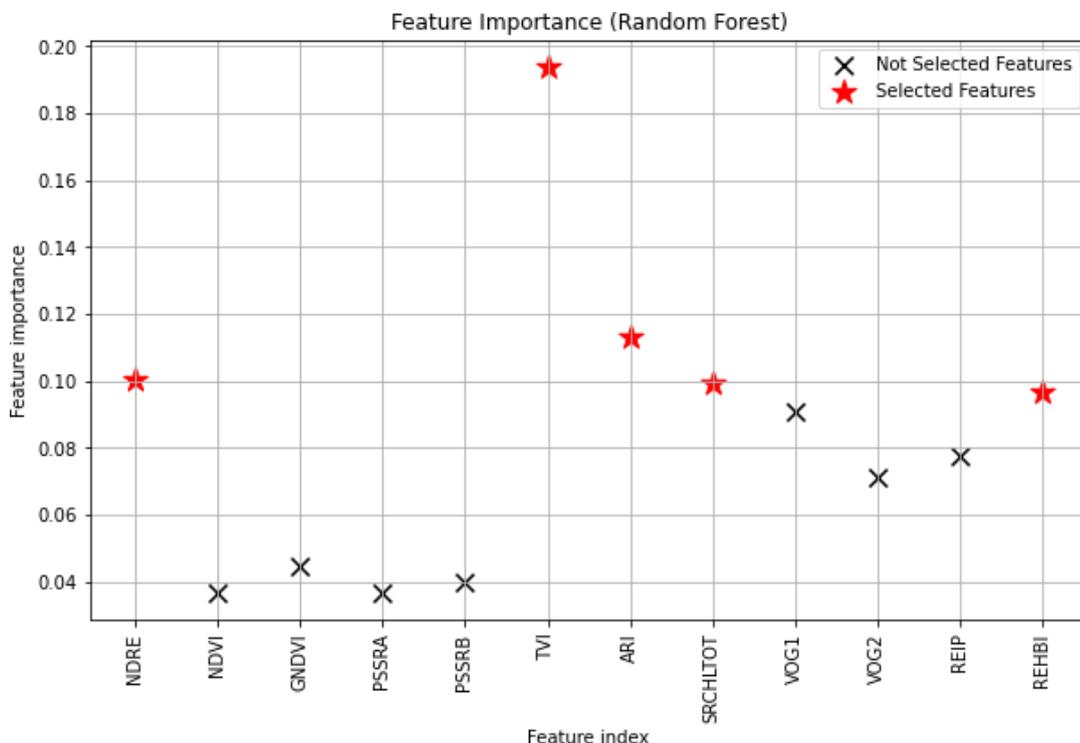
Αφού υπολογιστεί η σημαντικότητα όλων των χαρακτηριστικών, αυτά ταξινομούνται με βάση την τιμή της σημαντικότητάς τους, και επιλέγονται τα πέντε πιο σημαντικά. Η επιλογή αυτών των κορυφαίων χαρακτηριστικών έχει πολλαπλά οφέλη. Μειώνει την πολυπλοκότητα του μοντέλου, καθώς το μοντέλο εστιάζει μόνο στα πιο ουσιώδη χαρακτηριστικά. Βελτιώνει την απόδοση, αφού ένα μοντέλο με λιγότερα χαρακτηριστικά έχει μικρότερη πιθανότητα υπερπροσαρμογής. Επιπλέον, αυξάνει την ερμηνευσιμότητα του μοντέλου, καθώς οι σχέσεις μεταξύ των επιλεγμένων χαρακτηριστικών και της εξαρτημένης μεταβλητής γίνονται πιο κατανοητές.

Η χρήση του Random Forest για την εξαγωγή και επιλογή χαρακτηριστικών είναι ιδιαίτερα δημοφιλής στη μηχανική μάθηση, καθώς συνδυάζει υψηλή ακρίβεια και τη δυνατότητα ερμηνείας της σημαντικότητας των χαρακτηριστικών. Η επιλογή των πέντε καλύτερων χαρακτηριστικών καθιστά το μοντέλο πιο αποδοτικό, ενώ ταυτόχρονα επικεντρώνεται σε εκείνα τα στοιχεία που έχουν τη μεγαλύτερη επίδραση στις προβλέψεις. Στη συγκεκριμένη έρευνα χρησιμοποιείται η μέθοδος Entropy για τα δεδομένα που προέκυψαν από τις υπερφασματικές απεικονήσεις τόσο από τα φυτά αμπέλου όσο και από τα φυτά τομάτας των δύο πειραμάτων. Επιλέχθηκαν τα 5 κορυφαία χαρακτηριστικά για να χρησιμοποιηθούν ως είσοδοι στα μοντέλα μηχανικής μάθησης που αναπτύχθηκαν.

Τα αποτελέσματα από την επιλογή των χαρακτηριστικών για τα φυτά της αμπέλου και της τομάτας φαίνονται στις παρακάτω εικόνες (Εικόνα 1.5.1-18 και Εικόνα 1.5.1-19).



Εικόνα 1.5.1-18. Σημαντικότητα χαρακτηριστικών, όπως προέκυψαν από το μοντέλο Random Forest, για τα φασματικά δεδομένα που ελήφθησαν από τα φυτά της αμπέλου. Με κόκκινο αστέρι σημειώνονται τα 5 κορυφαία χαρακτηριστικά που έχουν επιλεγεί, ενώ με μαύρο 'x' σημειώνονται τα μη επιλεγμένα χαρακτηριστικά.



Εικόνα 1.5.1-19. Σημαντικότητα χαρακτηριστικών, όπως προέκυψαν από το μοντέλο Random Forest, για τα φασματικά δεδομένα που ελήφθησαν από τα φυτά τομάτας. Με κόκκινο αστέρι σημειώνονται τα 5 κορυφαία χαρακτηριστικά που έχουν επιλεγεί, ενώ με μαύρο 'x' εμφανίζονται τα μη επιλεγμένα χαρακτηριστικά.

Μέσω του αλγορίθμου αυτού, επιλέγονται οι σημαντικότεροι φασματικοί δείκτες, που συνεισφέρουν στην πρόβλεψη της κατάστασης της υγείας ή της παρουσίας ασθένειας στα φυτά.

Τα χαρακτηριστικά που επιλέχθηκαν ως πιο σημαντικά (με κόκκινο αστέρι στα διαγράμματα) περιλαμβάνουν διαφορετικούς φασματικούς δείκτες για τις δύο καλλιέργειες, με ορισμένα όμως να είναι κοινά. Έτσι, στην περίπτωση της καλλιέργειας του αμπελιού επιλέχθηκαν οι δείκτες TVI, ARI, SRCHLTOT, VOG1 και REHBI, ενώ στην τομάτα οι NDRE, TVI, ARI, SRCHLTOT και REHBI.

Τα σημαντικότερα χαρακτηριστικά στα φυτά αμπέλου περιλαμβάνουν, μεταξύ άλλων, τους δείκτες TVI, ARI, και SRCHLTOT, οι οποίοι σχετίζονται με την υγεία των φύλλων και την παρουσία χλωροφύλλης. Οι συγκεκριμένοι δείκτες είναι πιθανό να επιλέχθηκαν, καθώς είναι ευαίσθητοι σε αλλαγές που προκύπτουν λόγω ασθενειών, όπως το ωίδιο, που επηρεάζουν τη φωτοσυνθετική δραστηριότητα και τη βιοχημική σύνθεση των φύλλων.

Όσον αφορά στις τομάτες, οι δείκτες που επιλέχθηκαν, μεταξύ άλλων, περιλαμβάνουν τους NDRE, TVI, και REHBI. Ο δείκτης NDRE (Normalized Difference Red Edge) είναι γνωστός για την ευαισθησία του σε βιοτικές και αβιοτικές καταπονήσεις, ενώ ο REHBI, παρόλο που δημιουργήθηκε για την ανίχνευση της έντασης του *Fusarium Head Blight*, ως δείκτης σχετιζόμενος με την ένταση των ασθενειών, αντανακλά την επίδραση μυκητολογικών μολύνσεων, όπως είναι και η τεφρά σήψη.

Ένα ενδιαφέρον εύρημα είναι η παρουσία κοινών χαρακτηριστικών, όπως ο TVI (Triangular Vegetation Index) και ο ARI, στις δύο καλλιέργειες. Αυτό πιθανόν οφείλεται στη γενική ευαισθησία αυτών των δεικτών στις αλλαγές στη φυσιολογία των φυτών, που προκαλούνται από διάφορους τύπους καταπονήσεων. Πιο συγκεκριμένα, ο TVI ανιχνεύει τη χλωροφύλλη και τον γενικό βιοφυσικό status του φυτού, ενώ ο δείκτης ARI (Anthocyanin Reflectance Index) ανιχνεύει αλλαγές στα επίπεδα ανθοκυανινών, που συχνά σχετίζονται με αμυντικούς μηχανισμούς του φυτού σε ασθένειες ή καταπονήσεις.

Η διαφοροποίηση στη σημασία των δεικτών μεταξύ των δύο καλλιεργειών μπορεί να αποδοθεί στις διαφορετικές φυσιολογικές αντιδράσεις των δύο φυτών σε ασθένειες, οι οποίες καθορίζουν τις φασματικές τους υπογραφές. Για παράδειγμα, το ωίδιο στα αμπέλια και η τεφρά σήψη στις τομάτες προκαλούν διαφορετικές αλλαγές στα επίπεδα χλωροφύλλης και άλλων φυτικών χημικών. Επίσης, οι ασθένειες που επηρεάζουν τις δύο καλλιέργειες (ωίδιο και τεφρά σήψη) έχουν διαφορετικά παθογενετικά μονοπάτια, που επηρεάζουν τη φασματική απόκριση του φυτού σε διαφορετικές περιοχές του φάσματος.

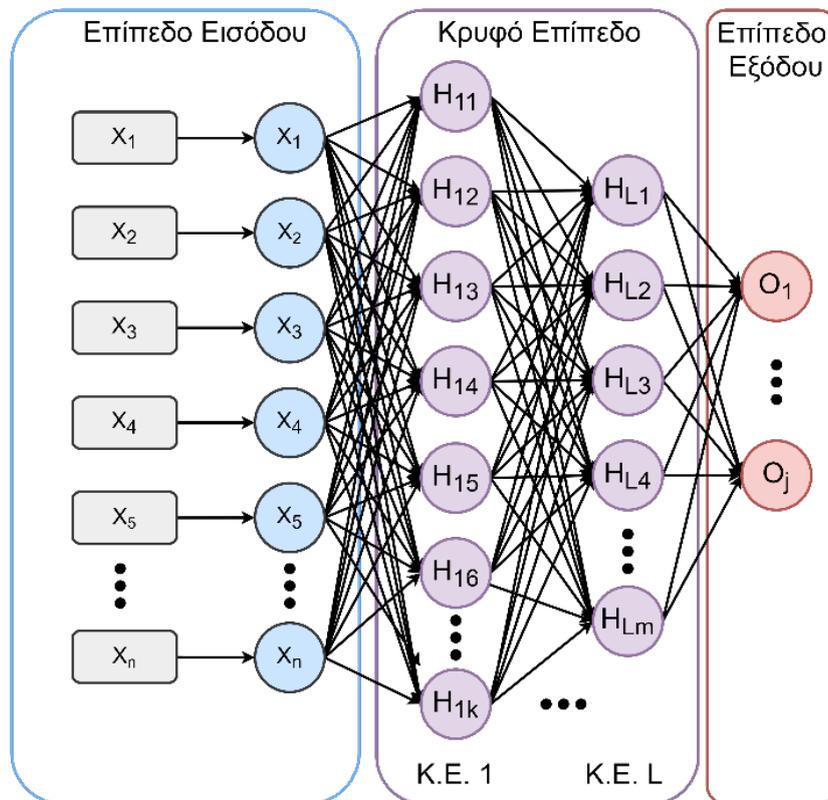
Η επιλογή των συγκεκριμένων χαρακτηριστικών από τον αλγόριθμο Random Forest βασίζεται στη δυνατότητά τους να εξηγούν τη μεταβλητότητα των δεδομένων. Τα χαρακτηριστικά που επελέγησαν παρουσιάζουν υψηλή συσχέτιση με την παρουσία ή την ένταση της ασθένειας. Επιπρόσθετα, είναι μη πλεονασματικά, δηλαδή προσφέρουν μοναδική πληροφορία που δεν μπορεί να αντικατασταθεί από άλλα χαρακτηριστικά,

ενώ επίσης συμβάλλουν στη μείωση του θορύβου και την αύξηση της ακρίβειας του μοντέλου.

Αλγόριθμοι Μηχανικής Μάθησης

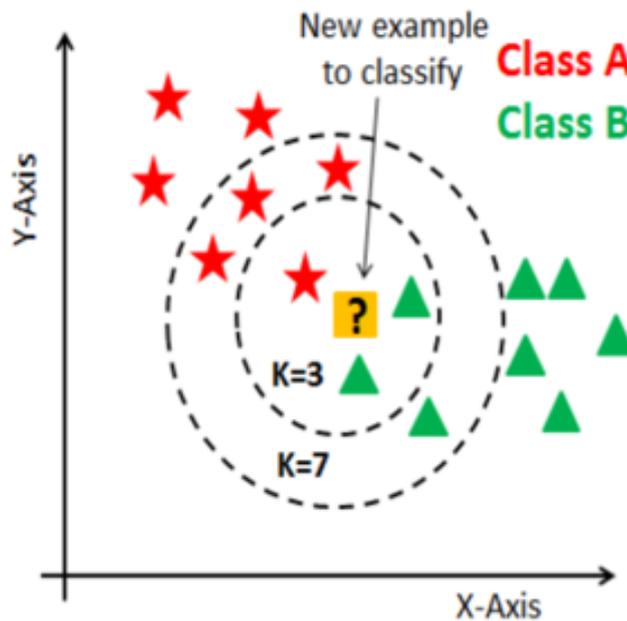
Η ανίχνευση καταπόνησης αποτελεί κρίσιμο πεδίο εφαρμογής της μηχανικής μάθησης, ειδικά σε τομείς όπως η μηχανολογία, η φυσική, η βιομηχανία και η ιατρική. Τα δεδομένα που συλλέγονται από αισθητήρες ή άλλες πηγές συχνά περιλαμβάνουν σύνθετα μοτίβα και αλληλεπιδράσεις, τα οποία μπορούν να αναλυθούν και να ερμηνευτούν με ακρίβεια μέσω σύγχρονων μεθόδων μηχανικής μάθησης. Τα μοντέλα **Multilayer Perceptron (MLP)**, **K-Nearest Neighbors (KNN)** και **Random Forest** διαδραματίζουν καίριο ρόλο σε αυτό το πλαίσιο, λόγω της ικανότητάς τους να επεξεργάζονται διαφορετικού τύπου δεδομένα και να εντοπίζουν σχέσεις που δεν είναι άμεσα εμφανείς.

Το Πολυστρωματικό Perceptron (Multilayer Perceptron, MLP) αποτελεί έναν τύπο τεχνητού νευρωνικού δικτύου που χρησιμοποιείται για την ανάλυση δεδομένων με σύνθετες σχέσεις. Αποτελείται από πλήρως συνδεδεμένα στρώματα, όπου κάθε νευρώνας συνδέεται με όλους τους νευρώνες του επόμενου στρώματος, επιτρέποντας τη διάχυση της πληροφορίας (Εικόνα 1.5.1-20). Η εισαγωγή μη γραμμικότητας μέσω συναρτήσεων ενεργοποίησης, όπως η ReLU, η σιγμοειδής ή η tanh, διευκολύνει την εκμάθηση σύνθετων μοτίβων. Η διαδικασία της οπισθοδιάδοσης (Backpropagation) επιτρέπει τη συνεχή βελτίωση των παραμέτρων του μοντέλου με βάση το σφάλμα στις προβλέψεις, καθιστώντας το MLP κατάλληλο για επιβλεπόμενη μάθηση. Στην ανίχνευση καταπόνησης, το MLP χρησιμοποιείται για την αναγνώριση πολύπλοκων προτύπων σε δεδομένα, όπως χρονικές σειρές δονήσεων, τάσεων ή θερμοκρασίας, καθώς και για την πρόβλεψη της μελλοντικής συμπεριφοράς συστημάτων βάσει ιστορικών δεδομένων.



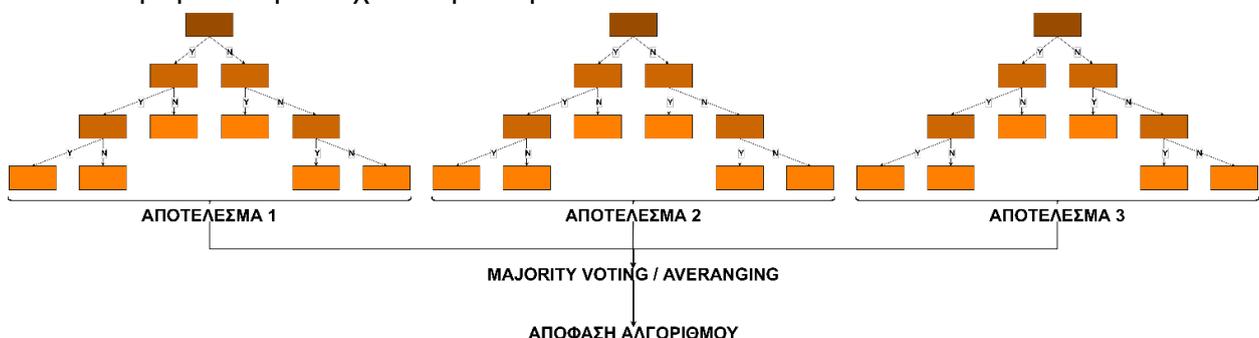
Εικόνα 1.5.1-20. Σχηματική αναπαράσταση της αρχιτεκτονικής του δικτύου multilayer perceptron (MLP), με ένα επίπεδο εισόδου, L κρυφά επίπεδα νευρώνων και ένα επίπεδο εξόδου.

Το K-Nearest Neighbors (KNN) είναι ένα μοντέλο μηχανικής μάθησης που βασίζεται στη μάθηση κατά περίπτωση. Δεν κατασκευάζει ρητό μοντέλο κατά την εκπαίδευση, αλλά βασίζεται στη γειτνίαση των δεδομένων για την ταξινόμηση ή την παλινδρόμηση. Το KNN χρησιμοποιεί τη μετρική απόστασης, όπως την ευκλείδεια ή τη Manhattan, για την εύρεση των πλησιέστερων γειτόνων ενός σημείου. Ο αριθμός των πλησιέστερων γειτόνων, που καθορίζεται από την τιμή K, επηρεάζει άμεσα την πρόβλεψη του μοντέλου (Εικόνα 1.5.1-21). Στην ανίχνευση καταπόνησης, το KNN χρησιμοποιείται για την ταξινόμηση καταστάσεων, όπως η "κανονική λειτουργία" και η "υπερβολική φόρτιση," καθώς και για την αναγνώριση ανωμαλιών ή περιοχών υψηλού κινδύνου καταπόνησης.



Εικόνα 1.5.1-21. Σχηματική αναπαράσταση της αρχιτεκτονικής του δικτύου K Nearest Neighbour, με 3 και 7 κοντινότερους γείτονες.

Το μοντέλο Random Forest βασίζεται στην τεχνική της μάθησης συνόλου (ensemble learning) και συνδυάζει πολλαπλά δέντρα απόφασης για την παροχή ακριβών και σταθερών προβλέψεων. Η διαδικασία του bagging (Bootstrap Aggregating) εξασφαλίζει τη διαφοροποίηση μεταξύ των δέντρων, μειώνοντας έτσι τον κίνδυνο υπερπροσαρμογής. Κάθε δέντρο εκπαιδεύεται σε τυχαία υποσύνολα των δεδομένων, ενώ η τελική πρόβλεψη προκύπτει από τη μέση πρόβλεψη (για παλινδρόμηση) ή την πλειοψηφική ψήφο (για ταξινόμηση) όλων των δέντρων (Εικόνα 1.5.1-22). Το Random Forest παρέχει επίσης πληροφορίες για τη σημασία των χαρακτηριστικών, βοηθώντας στην κατανόηση των παραγόντων που επηρεάζουν την καταπόνηση. Στην ανίχνευση καταπόνησης, το Random Forest χρησιμοποιείται για την ανάλυση δεδομένων από διάφορες πηγές, τον εντοπισμό σημαντικών παραγόντων που συνεισφέρουν στην καταπόνηση και την ανίχνευση ανωμαλιών.



Εικόνα 1.5.1-22. Σχηματική αναπαράσταση της αρχιτεκτονικής του αλγορίθμου μηχανικής μάθησης Random Forest, για 3 δέντρα απόφασης.

Συνολικά, το κάθε μοντέλο προσφέρει μοναδικά πλεονεκτήματα στην ανίχνευση καταπόνησης, ανάλογα με τη φύση των δεδομένων και τις απαιτήσεις της εφαρμογής.

Το MLP είναι κατάλληλο για την ανάλυση σύνθετων σχέσεων, το KNN αποδεικνύεται αποτελεσματικό για ταξινομήσεις βασισμένες σε τοπικά μοτίβα, ενώ το Random Forest παρέχει ανθεκτικότητα και δυνατότητα ερμηνείας των χαρακτηριστικών. Η επιλογή ή ο συνδυασμός αυτών των μοντέλων μπορεί να οδηγήσει σε σημαντικές βελτιώσεις στην ανίχνευση και την πρόληψη καταπόνησης, συμβάλλοντας στη σταθερότητα και την ασφάλεια των συστημάτων.

Επιλογή Υπερπαραμέτρων

Η επιλογή υπερπαραμέτρων για καθένα από τα χρησιμοποιούμενα μοντέλα θεωρείται ένα κρίσιμο βήμα στη μηχανική μάθηση, που έχει σημαντικό αντίκτυπο στην απόδοση κάθε μοντέλου που χρησιμοποιείται, όπως ο χρόνος εκπαίδευσης, η ακρίβεια του μοντέλου και η ικανότητά του να γενικεύεται σε ένα νέο σύνολο δεδομένων. Ο συντονισμός των υπερπαραμέτρων μπορεί να εφαρμοστεί χρησιμοποιώντας μια ποικιλία τεχνικών. Οι πιο κοινές μεταξύ αυτών των τεχνικών είναι η αναζήτηση πλέγματος και η βελτιστοποίηση Bayes. Στην παρούσα μελέτη, επιλέχθηκε η μέθοδος αναζήτησης πλέγματος. Αυτή η μέθοδος περιλαμβάνει τον έλεγχο μιας σειράς υπερπαραμέτρων και την επιλογή αυτής που δίνει την καλύτερη απόδοση σε ένα σύνολο επικύρωσης.

Για την εφαρμογή της μεθόδου αναζήτησης πλέγματος, δημιουργείται ένα πλέγμα υπερπαραμέτρων και στη συνέχεια μεθοδικά επεξεργάζονται και αξιολογούνται διάφοροι συνδυασμοί ακολουθώντας ένα προκαθορισμένο κριτήριο απόδοσης. Για να διασφαλιστεί ότι η επιτυχία του μοντέλου δεν εξαρτάται αποκλειστικά από τον τρόπο κατανομής των δεδομένων, χρησιμοποιούνται τεχνικές διασταυρούμενης επικύρωσης. Η διαδικασία επαναλαμβάνεται μέχρι να αξιολογηθεί κάθε πιθανός συνδυασμός στο πλέγμα. Στη συνέχεια, η ιδανική συλλογή υπερπαραμέτρων για το μοντέλο προσδιορίζεται μέσω του συνδυασμού των παραμέτρων με τρόπο που να αποδίδει το καλύτερο αποτέλεσμα όπως προσδιορίζεται από τη μέτρηση απόδοσης στο σύνολο επικύρωσης. Οι διαφορετικές υπερπαραμέτροι που εξετάστηκαν για καθένα από τα μοντέλα περιγράφονται στον (Πίνακας 1.5.1-2).

Πίνακας 1.5.1-2. Κατάλογος των πιθανών τιμών συντονισμού των υπερπαραμέτρων που χρησιμοποιούνται στη βελτιστοποίηση αναζήτησης πλέγματος για τα μοντέλα μηχανικής μάθησης random forest, k nearest neighbour(KNN) και πολυστρωματικών perceptron (MLP) που χρησιμοποιούνται σε αυτή τη μελέτη.

Μοντέλο	Υπερπαραμέτρος	Σύντομη Περιγραφή	Πιθανές Τιμές
Random Forest	Estimator count	Ελέγχει τον αριθμό των δέντρων απόφασης στο Random Forest.	5, 10, 15, 20, 30, 40, 50, 70, 90, 100, 120
	Criterion	Καθορίζει τη συνάρτηση που χρησιμοποιείται για τη μέτρηση της ποιότητας ενός διαχωρισμού σε κάθε κόμβο.	Mean Squared Error (MSE), Entropy
R	Max depth	Ελέγχει το μέγιστο βάθος κάθε δέντρου. Μεγάλο βάθος οδηγεί	2, 3, 4, 5, 6, 7

		σε υπερπροσαρμογή.	
	Min samples split	Καθορίζει τον ελάχιστο αριθμό δειγμάτων που απαιτούνται για τη διάσπαση ενός εσωτερικού κόμβου.	2, 3, 4, 5, 6
	Min samples leaf	Ελέγχει τον ελάχιστο αριθμό δειγμάτων που απαιτούνται για έναν κόμβο φύλλου. Μεγαλύτερη τιμή αποτρέπει την υπερπροσαρμογή.	1, 2, 3, 4, 5
KNN	n_neighbors	Ελέγχει τον αριθμό των πλησιέστερων γειτόνων. Επηρεάζει το bias-variance.	2, 3, 4, 5, 6, 7, 8, 9, 10
	weights	Καθορίζει πώς ζυγίζονται οι γείτονες κατά την ψήφο ή τον υπολογισμό του μέσου όρου.	uniform: Equal weights distance: Closer neighbors have higher influence
	p	Καθορίζει την παράμετρο δύναμης για την απόσταση Minkowski.	1: Manhattan distance 2: Euclidean distance
	metric	Ορίζει τη μετρική απόστασης για τον υπολογισμό των πλησιέστερων γειτόνων.	euclidean, manhattan, minkowski
MLP	Activation	Ελέγχει τη συνάρτηση ενεργοποίησης που χρησιμοποιείται στο MLP.	logistic, tanh, relu
	Solver	Ελέγχει τον αλγόριθμο βελτιστοποίησης για την εκπαίδευση των βαρών.	sgd, adam
	Hidden layer size	Ρυθμίζει τον αριθμό των στρωμάτων και τον αριθμό των κόμβων σε κάθε επίπεδο του MLP.	5 έως 180, με βήμα 5 για καθένα από τα 2 επίπεδα
	Learning rate	Ρυθμίζει το μέγεθος βήματος κατά την εκπαίδευση για τη σύγκλιση..	10^4 , 3×10^{-4} , 6×10^{-4} , 9×10^{-4} , 10^{-3} , 3×10^{-3} , 6×10^{-3} , 9×10^{-3} , 10^{-2} , 5×10^{-2} , 10^{-1}
	Learning rate state	Ρυθμίζει τον τρόπο μεταβολής του Learning Rate (LR) στον αλγόριθμο SGD.	constant, invscaling, adaptive

Στον Πίνακα 1.5.1-1, οι σημειωμένες συναρτήσεις ενεργοποίησης είναι η λογιστική, η υπερβολική εφαπτομένη (tanh) και η διορθωμένη γραμμική μονάδα (relu). Επιπλέον, ο βελτιστοποιητής επίλυσης θα μπορούσε να ήταν είτε η στοχαστική κατάβαση δυναμικού (sgd) είτε η προσαρμοστική εκτίμηση ροπής (adam). Τέλος, ο προγραμματισμός του ρυθμού μάθησης επιλέχθηκε είτε να είναι σταθερός είτε με αντίστροφη κλίμακα (invscaling). Το τελευταίο είναι ικανό να μειώνει σταδιακά το LR σε κάθε επανάληψη και προσαρμογή. Με αυτόν τον τρόπο το LR παραμένει σταθερό έως ότου το σφάλμα

εκπαίδευσης σταματήσει να μειώνεται, όπου διαιρείται με το 5.

Τα αποτελέσματα από την επιλογή των υπερπαραμέτρων για το κάθε μοντέλο και το κάθε πείραμα φαίνεται στον Πίνακα 1.5.1-3.

Πίνακας 1.5.1-3. Επιλεγμένες τιμές υπερπαραμέτρων, όπως προέκυψαν μετά από τη χρήση του αλγορίθμου αναζήτησης πλέγματος, για το πείραμα των φυτών αμπέλου και τομάτας, για καθένα από τα μοντέλα μηχανικής μάθησης που χρησιμοποιήθηκαν.

Μοντέλο	Υπερπαραμέτρος	Πείραμα Αμπέλου	Πείραμα Τομάτας
Random Forest	Estimator count	20	52
	Criterion	Entropy	MSE
	Max depth	15	32
	Min samples split	2	3
	Min samples leaf	1	2
KNN	n_neighbors	3	8
	weights	uniform	uniform
	p	-	-
	metric	euclidean	minkowski
MLP	Activation	relu	tanh
	Solver	adam	adam
	Hidden layer size	(60, 100)	(15, 80)
	Learning rate	0.0001	0.009
	Learning rate state	constant	constant

Μετρικές Αξιολόγησης των μοντέλων

Τα δεδομένα που παρουσιάζονται στην παρούσα μελέτη χωρίστηκαν τυχαία σε δεδομένα εκπαίδευσης και δεδομένα προβλέψεων. Ο χωρισμός πραγματοποιήθηκε βάση την συνήθη αναλογία 70-30% για τα δεδομένα εκπαίδευσης-επαλήθευσης αντίστοιχα.

Τα αποτελέσματα των επιδόσεων των ταξινομητών που χρησιμοποιήθηκαν απεικονίζονται με τη χρήση του πίνακα σύγχυσης (confusion matrix), που είναι ένας δισδιάστατος πίνακας, όπου οι στήλες αντιστοιχούν στις προβλέψεις και οι γραμμές στις πραγματικές τιμές της κλάσης που μελετάται

Στον παραπάνω πίνακα το πλήθος των επιτυχών προβλέψεων για την κλάση των θετικών παρατηρήσεων ονομάζεται πραγματικές θετικές προβλέψεις (True Positive - TP), ενώ το πλήθος των επιτυχών προβλέψεων για την κλάση των αρνητικών παρατηρήσεων ονομάζεται πραγματικές αρνητικές προβλέψεις (True Negative - TN). Από την άλλη το πλήθος των αποτυχημένων θετικών προβλέψεων ονομάζεται ψευδείς αρνητικές προβλέψεις (False Negative - FN) ενώ το αντίστοιχο πλήθος των αποτυχημένων αρνητικών προβλέψεων ονομάζεται ψευδείς θετικές προβλέψεις (False Positive - FP).

Για την αξιολόγηση των αποτελεσμάτων των πινάκων σύγχυσης χρησιμοποιήθηκαν οι μέθοδοι εκτίμησης της ακρίβειας accuracy, precision και F1 score. Οι αντίστοιχες εξισώσεις που προσδιορίζουν την καθεμία από τις μεθόδους α φαίνονται παρακάτω (Εξ. 1, Εξ. 2 και Εξ. 3).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (\text{Εξ. 1})$$

$$Precision = \frac{TP}{TP+FN} \quad (\text{Εξ. 2})$$

$$F1score = 2 \cdot \frac{Recall \cdot Precision}{Recall + Precision}, \quad \text{όπου: } Recall = \frac{TP}{TP+FN} \quad (\text{Εξ. 3})$$

Ανίχνευση ωιδίου σε καλλιέργεια αμπελιού

Τα αποτελέσματα για καθένα από τα μοντέλα που χρησιμοποιήθηκαν για την έγκαιρη διάγνωση του ωιδίου στα αμπέλια, για κάθε μία από τις τρεις κλάσεις που δημιουργήθηκαν, φαίνονται στον Πίνακας 1.5.1-4.

Πίνακας 1.5.1-4. Αποτελέσματα αλγορίθμων μηχανικής μάθησης (ML) κατά τη φάση εκπαίδευσης και δοκιμής για το πείραμα με τα φυτά αμπέλου. Οι τιμές με έντονους χαρακτήρες αντιπροσωπεύουν τα μοντέλα ML με την καλύτερη απόδοση.

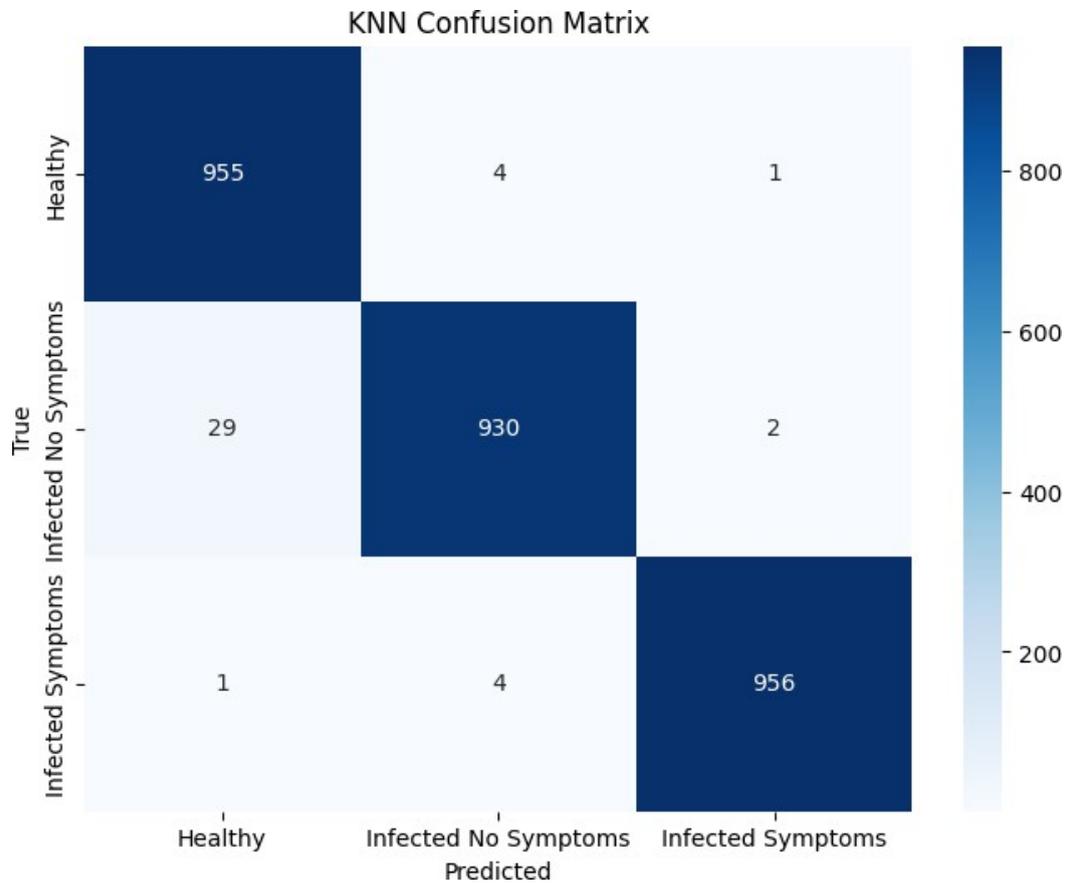
Μοντέλο	Metric	Εκπαίδευση (Training)	Έλεγχος (Test)
RF*	Accuracy	0.99	0.99
	Precision	1.00	0.99
	F1-Score	0.99	0.99
KNN*	Accuracy	0.99	0.99
	Precision	0.98	0.98
	F1-Score	0.99	0.99
MLP*	Accuracy	0.98	0.98
	Precision	0.99	0.99
	F1-Score	0.98	0.99

*RF: random forests, LS-SVM: least-squares support vector machine, MLP: multilayer perceptron.

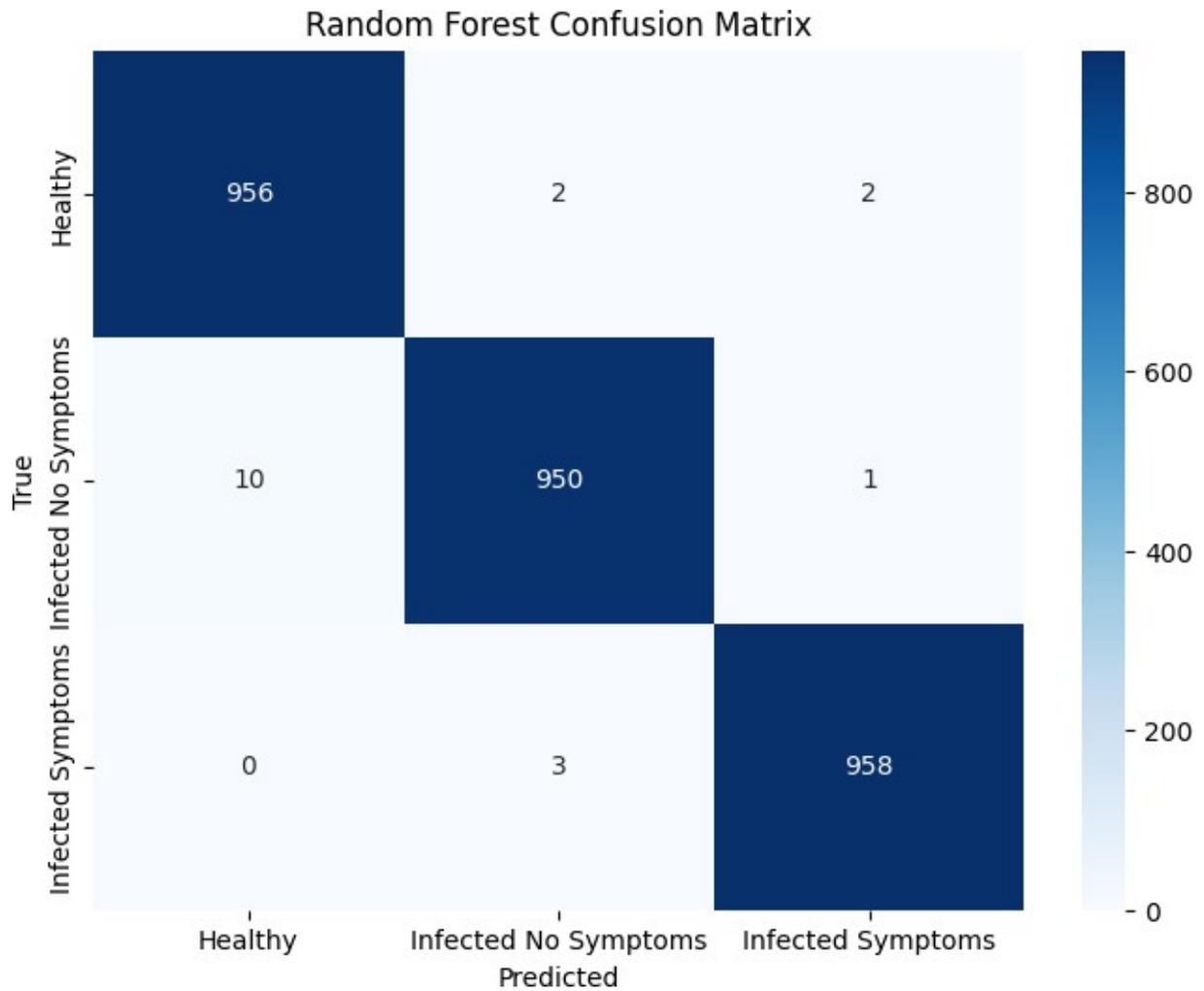
Όπως φαίνεται από τον Πίνακα 1.5.1-4, όλα τα μοντέλα παρουσιάζουν ακρίβεια 0.99, που είναι σχεδόν άριστο αποτέλεσμα. Επίσης, αξίζει να σημειωθεί ότι στη φάση της εκπαίδευσης η ακρίβεια ήταν είτε υψηλότερη είτε ίση με τη φάση της δοκιμής, πράγμα που υποδηλώνει ότι η εκμάθηση των δεδομένων ήταν εξαιρετικά αποδοτική. Το Random Forest σημείωσε Ακρίβεια (Precision) ίση με 1.00, γεγονός που σημαίνει ότι αναγνωρίζει σωστά όλες τις θετικές τιμές της κλάσης του κατά τη διαδικασία της εκπαίδευσης. Τέλος, οι τιμές των F1-Score σε όλα τα μοντέλα είναι σχεδόν άριστες και αυτό επιβεβαιώνει την ιδιαίτερα υψηλή δυναμική των μοντέλων που αναπτύχθηκαν όσο αφορά την ανίχνευση του ωιδίου σε προσυμπτωματικό στάδιο. Ταυτόχρονα, το γεγονός ότι οι τιμές των F1-Score ήταν τόσο κοντά με τα αποτελέσματα της ακρίβειας επιβεβαιώνει ότι υπήρχε ίση κατανομή των δεδομένων στις διαφορετικές κλάσεις που προέκυψαν μετά και από τη διαδικασία υποδειγματοληψίας των συλλεχθέντων φασματικών δεδομένων.

Τα αποτελέσματα του παραπάνω πίνακα για την φάση της δοκιμής του κάθε μοντέλου

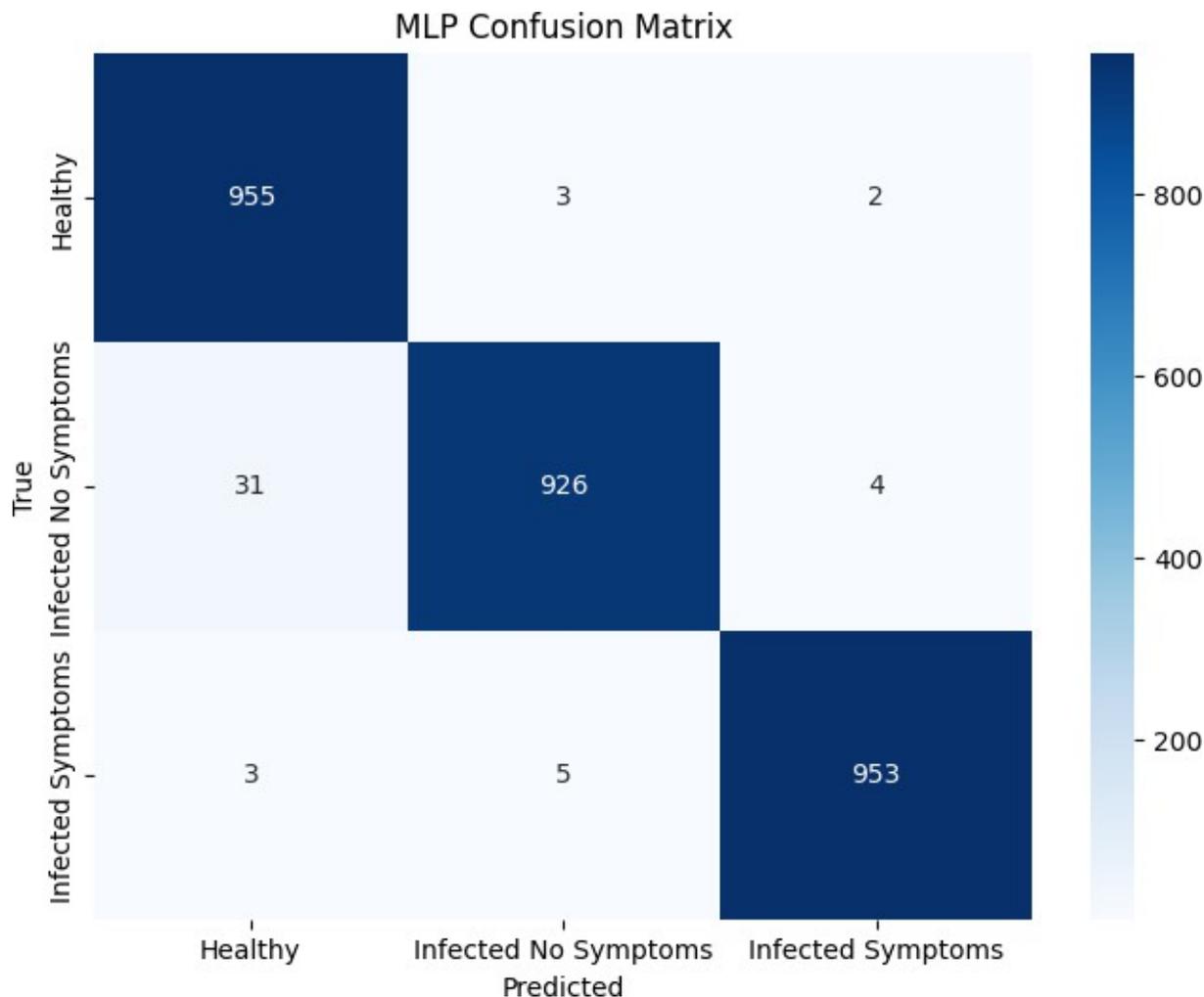
απεικονίζονται και στους πίνακες σύγχυσης που ακολουθούν (Εικόνα 1.5.1-23, Εικόνα 1.5.1-24 και Εικόνα 1.5.1-25).



Εικόνα 1.5.1-23. Πίνακας σύγχυσης των αποτελεσμάτων των φασματικών αποκρίσεων του αμπελιού με τη μέθοδο KNN.



Εικόνα 1.5.1-24. Πίνακας σύγχυσης των αποτελεσμάτων των φασματικών αποκρίσεων του αμπελιού με τη μέθοδο Random Forest.



Εικόνα 1.5.1-25. Πίνακας σύγχυσης των αποτελεσμάτων των φασματικών αποκρίσεων του αμπελιού με τη μέθοδο MLP.

Από τους πίνακες σύγχυσης των παραπάνω εικόνων, φαίνεται ότι όντως κυριαρχούν τα σωστά ταξινομημένα παραδείγματα, καθώς η κύρια διαγώνιος έχει σχεδόν το σύνολο του πλήθους των παραδειγμάτων. Υπάρχουν ελάχιστα λανθασμένα ταξινομημένες τιμές και ως επί το πλείστον αυτές είναι μεταξύ των πρώτων δύο κλάσεων, γεγονός που δεν αποτελεί έκπληξη, καθώς και οι δύο αυτές οι κλάσεις, οπτικά τουλάχιστον, φαίνονται να είναι υγιείς. Έτσι είναι πιθανό η ένταση της χλωροφύλλης να είναι υψηλή στο μολυσμένο φύλλο ή αντίστοιχα η ένταση της ασθένειας να είναι πολύ χαμηλή, καθώς το φύλλο δύναται να βρίσκεται στα πρώτα στάδια της μόλυνσης με την ασθένεια του ωιδίου. Τέλος, αξίζει να σημειωθεί ότι ελάχιστα παραδείγματα της ασθενούς κλάσης φαίνεται από τους πίνακες σύγχυσης να είναι λάθος ταξινομημένες.

Ανίχνευση Τεφράς Σήψης σε Τομάτες

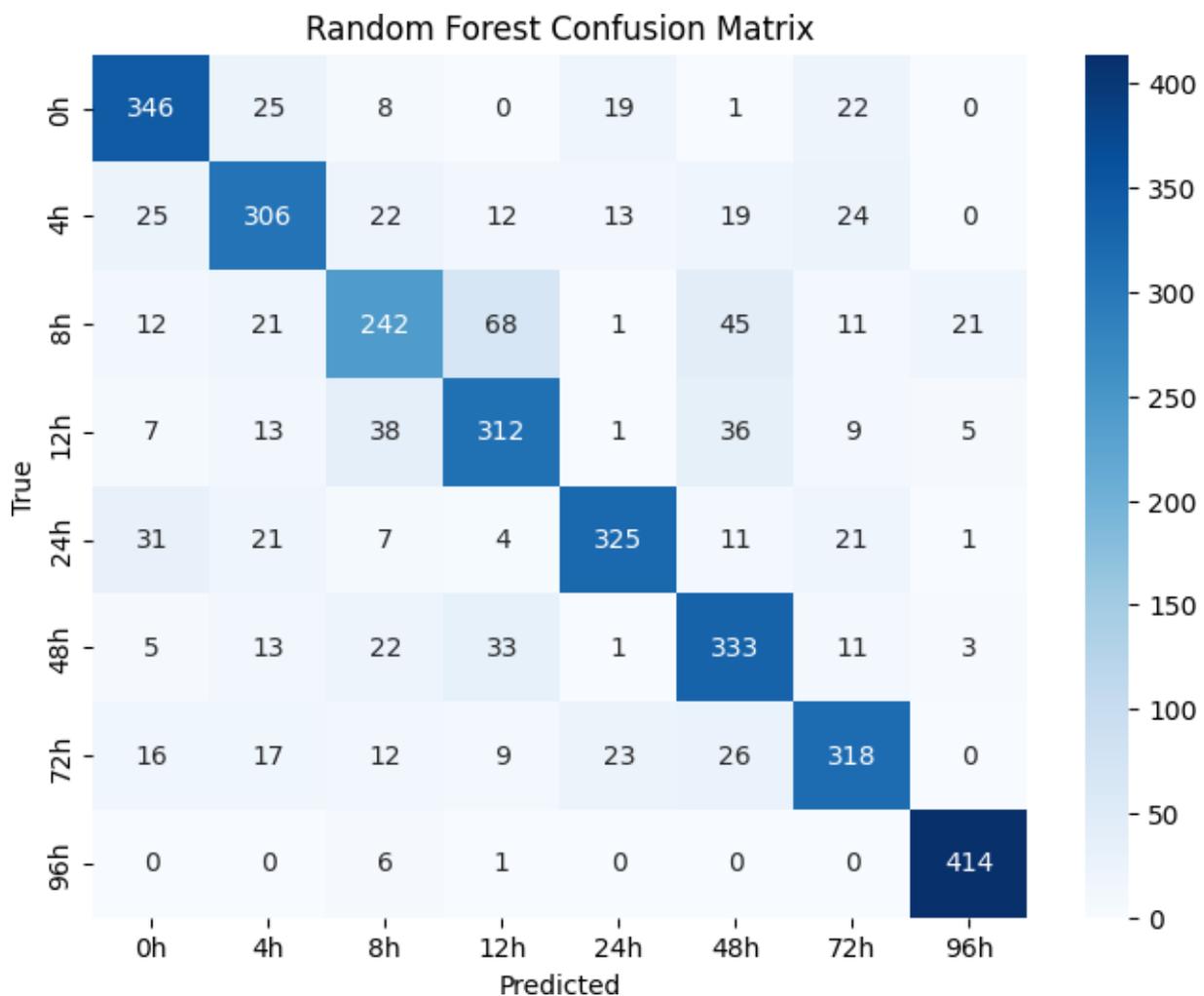
Πίνακας 1.5.1-5. Αποτελέσματα αλγορίθμων μηχανικής μάθησης (ML) κατά τη φάση εκπαίδευσης και δοκιμής για το πείραμα με τα φυτά τομάτας. Οι τιμές με έντονους χαρακτήρες αντιπροσωπεύουν τα μοντέλα ML με την καλύτερη απόδοση.

Model	Metric	Εκπαίδευση (Training)	Έλεγχος (Test)
RF*	Accuracy	0.79	0.77
	Precision	0.80	0.77
	F1-Score	0.78	0.77
KNN*	Accuracy	0.60	0.59
	Precision	0.60	0.59
	F1-Score	0.59	0.59
MLP*	Accuracy	0.84	0.81
	Precision	0.86	0.82
	F1-Score	0.83	0.81

Ο Πίνακας 1.5.1-5 παρουσιάζει τα αποτελέσματα τριών διαφορετικών αλγορίθμων μηχανικής μάθησης (Random Forest - RF, k-Nearest Neighbors - KNN, Multilayer Perceptron - MLP) που εφαρμόστηκαν για την ανάλυση δεδομένων από το πείραμα με τα φυτά τομάτας. Από τα αποτελέσματα διακρίνεται ότι το MLP εμφανίζει την καλύτερη συνολική απόδοση, με ακρίβεια 0.84 στη φάση εκπαίδευσης και 0.81 στη φάση ελέγχου, γεγονός που υποδεικνύει εξαιρετική ικανότητα γενίκευσης. Επιπλέον, οι τιμές Ακρίβειας (Precision) και F1-Score για το MLP είναι 0.86 και 0.83 αντίστοιχα στην εκπαίδευση, ενώ ελαφρώς μειώνονται σε 0.82 και 0.81 στη δοκιμή, παραμένοντας ωστόσο σε υψηλά επίπεδα.

Το μοντέλο Random Forest (RF) εμφανίζει καλή αλλά υποδεέστερη απόδοση σε σχέση με το MLP, με ακρίβεια 0.79 στη φάση εκπαίδευσης και 0.77 στη δοκιμή. Το Precision και το F1-Score του RF ανέρχονται σε 0.80 και 0.78 αντίστοιχα στην εκπαίδευση, ενώ στη δοκιμή φτάνουν στο 0.77, γεγονός που δείχνει σταθερή αλλά μέτρια απόδοση σε σύγκριση με το MLP. Από την άλλη πλευρά, το μοντέλο k-Nearest Neighbors (KNN) παρουσιάζει τη χαμηλότερη απόδοση από τα τρία μοντέλα, με ακρίβεια 0.60 στη φάση εκπαίδευσης και 0.59 στη δοκιμή. Αντίστοιχα, οι τιμές Precision και F1-Score για το KNN είναι 0.60 και 0.59, αντίστοιχα, υποδηλώνοντας την περιορισμένη δυνατότητά του να διακρίνει τις κλάσεις των δεδομένων.

Συνολικά, το MLP αναδεικνύεται ως η καλύτερη επιλογή για την ανάλυση των δεδομένων του πειράματος, καθώς συνδυάζει την υψηλότερη ακρίβεια, precision, και F1-Score. Το RF αποτελεί μια αξιόπιστη εναλλακτική λύση, παρόλο που η απόδοσή του είναι χαμηλότερη, ενώ το KNN αποδεικνύεται λιγότερο κατάλληλο για το συγκεκριμένο πρόβλημα, παρουσιάζοντας σημαντικά χαμηλότερες τιμές στις μετρικές αξιολόγησης. Παρακάτω φαίνονται τα αποτελέσματα των μοντέλων ανά κλάση, όπως προκύπτουν από τους πίνακες σύγχυσης που ακολουθούν (Εικόνα 1.5.1-26, Εικόνα 1.5.1-28, Εικόνα 1.5.1-30).



Εικόνα 1.5.1-26. Πίνακας σύγχυσης των αποτελεσμάτων των φασματικών αποκρίσεων των φυτών της τομάτας με τη μέθοδο RF.

Ο πίνακας σύγχυσης (Εικόνα 1.5.1-26) παρουσιάζει την απόδοση του μοντέλου Random Forest (RF) για την ταξινόμηση των φασματικών αποκρίσεων φυτών τομάτας σε διαφορετικά χρονικά διαστήματα. Ο πίνακας δείχνει τις σωστές και λανθασμένες προβλέψεις του μοντέλου, με τις γραμμές να αντιπροσωπεύουν τις πραγματικές κατηγορίες και τις στήλες τις προβλεφθείσες κατηγορίες.

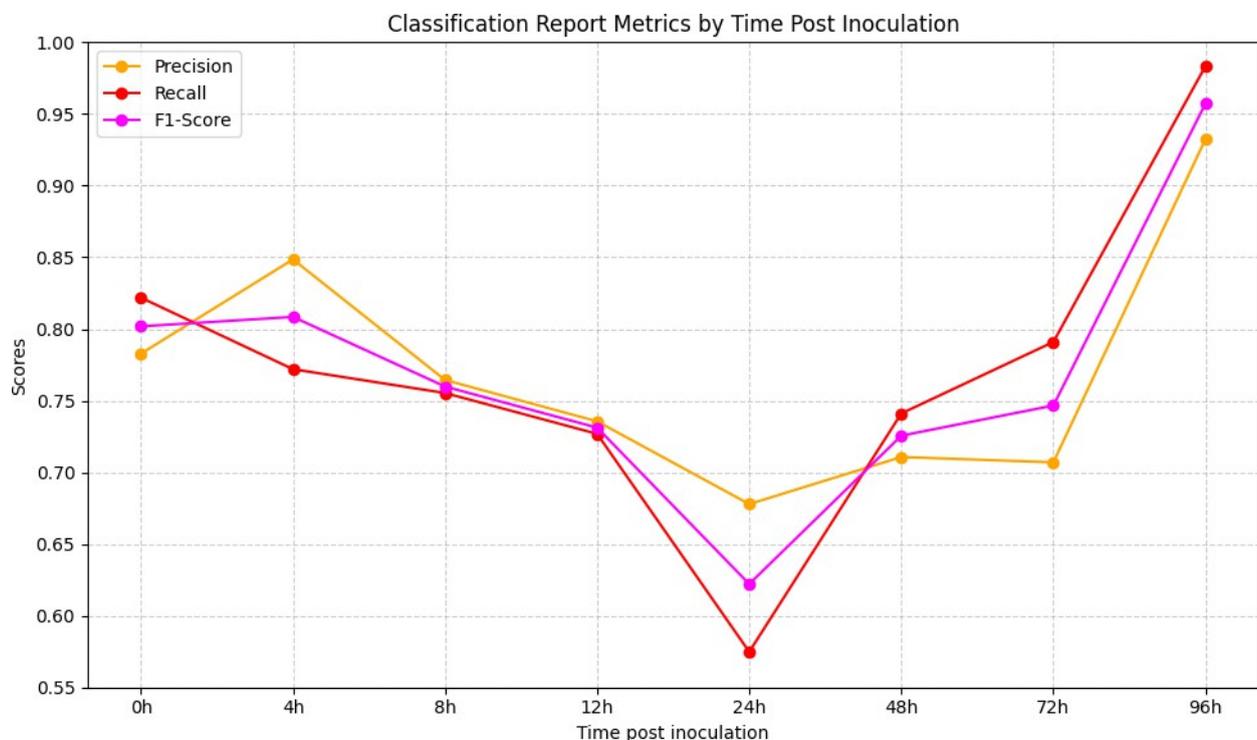
Η κατηγορία "96h" εμφανίζει υψηλή απόδοση, με το μοντέλο να προβλέπει σωστά 414

δείγματα, γεγονός που αποδίδεται στην έντονη αλλαγή της ανακλαστικής ικανότητας των φυτών σε αυτό το στάδιο λόγω της εμφάνισης των συμπτωμάτων. Παρότι η κατηγορία αυτή έχει υψηλή ακρίβεια, υπάρχουν ορισμένες λανθασμένες προβλέψεις που κατηγοριοποιήθηκαν ως "96h" ή αντίστροφα. Η κατηγορία "0h" σημειώνει επίσης καλή απόδοση με 346 σωστές προβλέψεις, αν και παρατηρούνται λανθασμένες ταξινομήσεις σε κοντινές κατηγορίες, όπως "4h" και "8h".

Οι κατηγορίες "8h" και "12h" εμφανίζουν μεγαλύτερη δυσκολία στη διάκριση, καθώς το μοντέλο μπερδέψε σημαντικό αριθμό δειγμάτων μεταξύ αυτών των δύο κατηγοριών, αλλά και με άλλες κοντινές χρονικές κατηγορίες. Για παράδειγμα, η κατηγορία "8h" είχε 242 σωστές προβλέψεις, αλλά 68 δείγματα ταξινομήθηκαν λανθασμένα ως "12h". Παρόμοια, η κατηγορία "12h" είχε 312 σωστές προβλέψεις, αλλά υπήρξαν 38 δείγματα που ταξινομήθηκαν ως "8h" και 36 ως "24h". Αυτό υποδεικνύει ότι οι φασματικές αποκρίσεις σε αυτά τα ενδιάμεσα χρονικά διαστήματα είναι πιο δύσκολο να διακριθούν.

Η κατηγορία "48h" σημειώνει καλή απόδοση με 333 σωστές προβλέψεις, ενώ η κατηγορία "72h" είχε 318 σωστές προβλέψεις, αν και στις δύο περιπτώσεις υπήρξαν κάποιες συγχύσεις με γειτονικές κατηγορίες, όπως η "24h" ή η "96h".

Συνολικά, το Random Forest παρουσιάζει υψηλή ακρίβεια, ιδιαίτερα σε κατηγορίες όπου οι φασματικές αποκρίσεις διαφοροποιούνται έντονα, όπως στις "96h" και "0h". Ωστόσο, στις ενδιάμεσες χρονικές κατηγορίες παρατηρείται μεγαλύτερη σύγχυση, ειδικά μεταξύ χρονικά κοντινών κατηγοριών όπως οι "8h", "12h" και "24h". Αυτό μπορεί να οφείλεται σε μικρότερες διαφορές στα φασματικά χαρακτηριστικά σε αυτά τα στάδια.



Εικόνα 1.5.1-27. Μετρικές ταξινόμησης για κάθε μία από τις κλάσεις που προέκυψαν από το πείραμα με τα φυτά της τομάτας για τον αλγόριθμο RF.

Ο πίνακας σύγχυσης (Εικόνα 1.5.1-26) και το διάγραμμα των μετρικών Precision, Recall και F1-Score (Εικόνα 1.5.1-27) αποτυπώνουν την απόδοση του Random Forest (RF) στην ταξινόμηση φασματικών αποκρίσεων φυτών τομάτας. Η κατηγορία "96h" εμφανίζει εξαιρετική απόδοση και αυτό φαίνεται και στο διάγραμμα, όπου οι μετρικές Precision, Recall και F1-Score για την κατηγορία αυτή είναι πολύ υψηλές, αγγίζοντας σχεδόν το

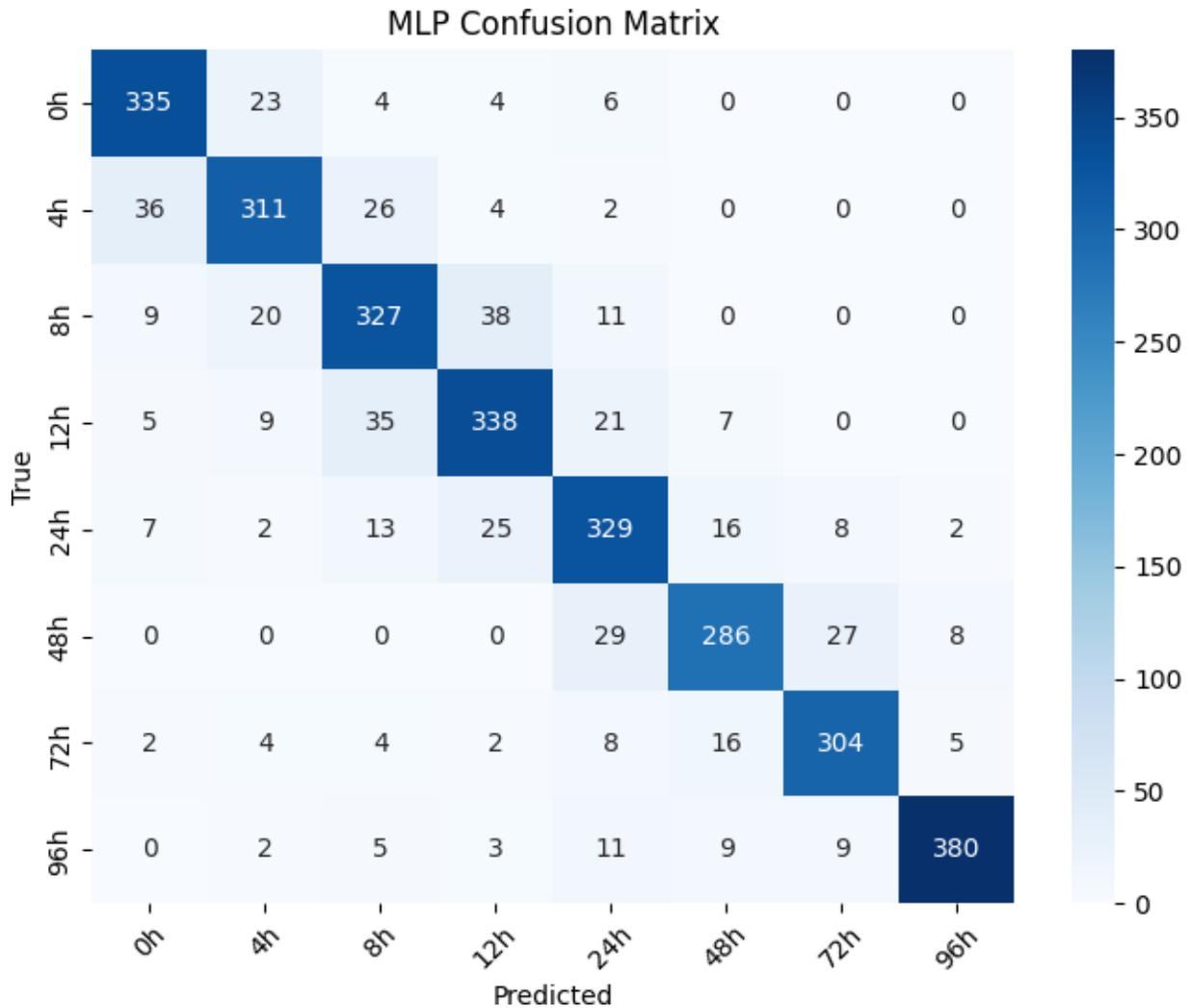
1. Η απόδοση αυτή δικαιολογείται από την έντονη διαφοροποίηση στα φασματικά χαρακτηριστικά της συγκεκριμένης κατηγορίας, που καθιστά τη διάκριση ευκολότερη.

Αντίστοιχα η σταθερή απόδοση στην κατηγορία "0h", όπου σύμφωνα με τον πίνακα σύγχυσης παρουσιάζει 346 σωστές προβλέψεις, με περιορισμένες λανθασμένες ταξινομήσεις προς τις κατηγορίες "4h" και "8h", αντικατοπτρίζεται και στο διάγραμμα της Εικόνα 1.5.1-27, όπου οι μετρικές είναι υψηλές, αν και όχι απόλυτες. Στην κατηγορία "4h", παρατηρούνται 306 σωστές προβλέψεις, με ορισμένες συγχύσεις με τις κατηγορίες "0h" και "8h", κάτι που επηρεάζει τις αντίστοιχες τιμές Recall και Precision, όπως φαίνεται στο διάγραμμα.

Οι κατηγορίες "8h" και "12h" εμφανίζουν τις περισσότερες προκλήσεις για το μοντέλο. Η κατηγορία "8h" σημειώνει 242 σωστές προβλέψεις, με αρκετές λανθασμένες ταξινομήσεις προς τις κατηγορίες "12h" και "24h". Παρόμοια, η κατηγορία "12h" έχει 312 σωστές προβλέψεις, αλλά με συγχύσεις προς "8h" και "24h". Οι χαμηλότερες τιμές στις μετρικές Precision, Recall και F1-Score για αυτές τις κατηγορίες στο διάγραμμα υποδεικνύουν την αυξημένη πολυπλοκότητα στη διάκριση των χαρακτηριστικών τους.

Η κατηγορία "24h" παρουσιάζει 325 σωστές προβλέψεις, αλλά υπήρξαν λανθασμένες ταξινομήσεις κυρίως προς τις κατηγορίες "12h" και "48h". Παρόλο που η απόδοση της κατηγορίας είναι καλύτερη από τις ενδιάμεσες φάσεις, οι μετρικές στο διάγραμμα παραμένουν μέτριες. Αντίθετα, οι κατηγορίες "48h" και "72h" σημειώνουν βελτιωμένη απόδοση, με 333 και 318 σωστές προβλέψεις αντίστοιχα. Οι τιμές Precision, Recall και F1-Score για αυτές τις κατηγορίες στο διάγραμμα δείχνουν ανοδική τάση, καθώς οι φασματικές αποκρίσεις σταδιακά διαφοροποιούνται περισσότερο.

Συνολικά, το Random Forest αποδεικνύει την αποτελεσματικότητά του στις κατηγορίες με έντονες διαφοροποιήσεις, όπως η "96h", ενώ δυσκολεύεται στις ενδιάμεσες φάσεις, όπως οι "8h" και "12h". Η καλή απόδοση στις αρχικές και μεταγενέστερες κατηγορίες ενισχύει τη χρησιμότητα του μοντέλου, αλλά τα ευρήματα υποδεικνύουν περιθώρια βελτίωσης στις φάσεις όπου τα χαρακτηριστικά αλληλεπικαλύπτονται. Το διάγραμμα συμπληρώνει την εικόνα του πίνακα σύγχυσης, επισημαίνοντας τις περιοχές όπου το μοντέλο αποδίδει καλύτερα ή έχει προκλήσεις.



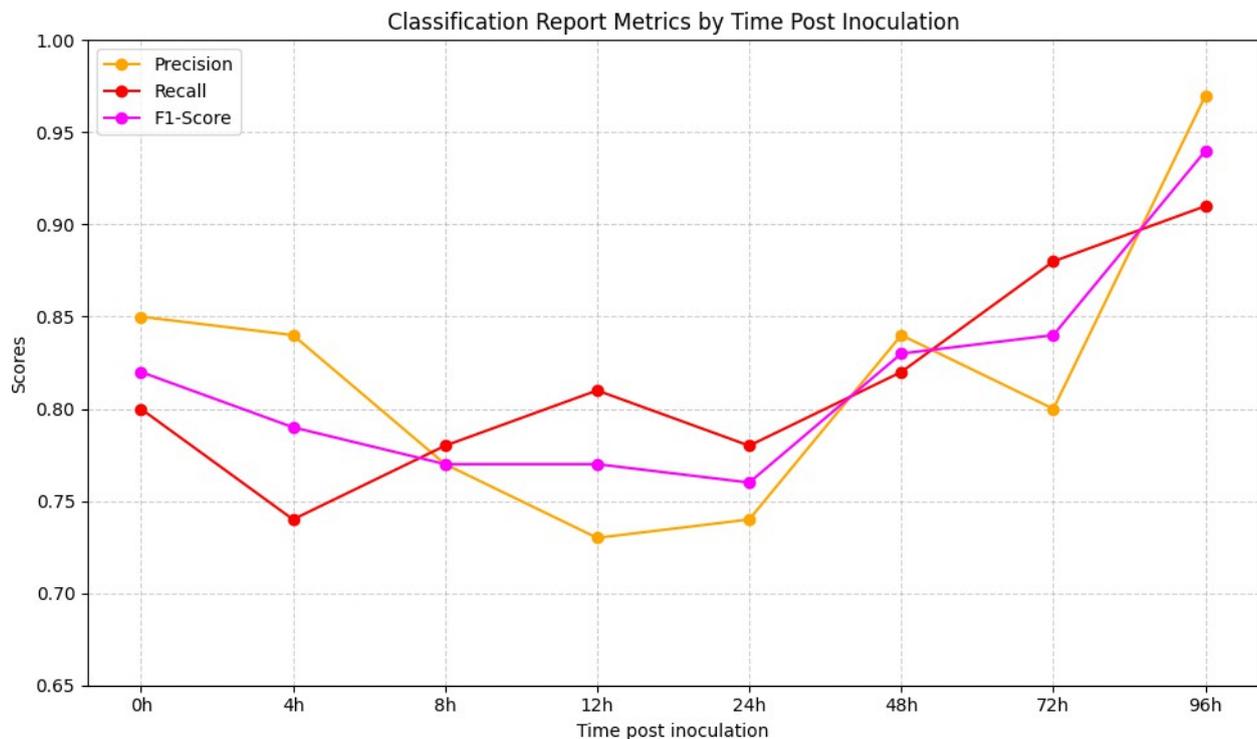
Εικόνα 1.5.1-28. Πίνακας σύγχυσης των αποτελεσμάτων των φασματικών αποκρίσεων των φυτών της τομάτας με τη μέθοδο MLP.

Ο πίνακας σύγχυσης (Εικόνα 1.5.1-28) παρουσιάζει την απόδοση του μοντέλου MLP για την ταξινόμηση των φασματικών αποκρίσεων φυτών τομάτας σε διαφορετικά χρονικά διαστήματα. Και σε αυτόν τον πίνακα, η κλάση των "96h" σημειώνει εξαιρετική απόδοση με 380 σωστές προβλέψεις, γεγονός που δικαιολογείται από τη σημαντική αλλαγή στη φασματική απόκριση σε αυτό το στάδιο λόγω της έντονης εμφάνισης των συμπτωμάτων. Παρόλα αυτά, υπήρξαν κάποιες ελάχιστες συγχύσεις, με 9 δείγματα να ταξινομούνται λανθασμένα σε γειτονικές κατηγορίες, όπως η "72h".

Η κατηγορία "0h" είχε 335 σωστές προβλέψεις, αλλά παρατηρήθηκαν λανθασμένες ταξινομήσεις, κυρίως προς την κατηγορία "4h" (23 δείγματα) και την "8h" (4 δείγματα). Παρά την καλή διακριτική της ικανότητα, κάποιες συγχύσεις με κοντινές χρονικές κατηγορίες υποδεικνύουν παρόμοια φασματικά χαρακτηριστικά. Οι κατηγορίες "8h" και "12h" παρουσίασαν μεγαλύτερες δυσκολίες στη διάκριση. Η κατηγορία "8h" είχε 327 σωστές προβλέψεις, αλλά 20 δείγματα ταξινομήθηκαν ως "4h" και 38 ως "12h". Αντίστοιχα, η κατηγορία "12h" είχε 338 σωστές προβλέψεις, με λανθασμένες ταξινομήσεις κυρίως προς τις "8h" (35 δείγματα) και "24h" (21 δείγματα). Αυτές οι συγχύσεις αναδεικνύουν τη δυσκολία διάκρισης μεταξύ χρονικά κοντινών σταδίων.

Η κατηγορία "24h" είχε 329 σωστές προβλέψεις, με μικρές συγχύσεις κυρίως προς τις "12h" (25 δείγματα) και "48h" (16 δείγματα). Παρόμοια, η κατηγορία "48h" παρουσίασε 286 σωστές προβλέψεις, αλλά υπήρχαν συγχύσεις με τις κατηγορίες "24h" (29 δείγματα) και "72h" (27 δείγματα). Η κατηγορία "72h" είχε 304 σωστές προβλέψεις, με ορισμένα δείγματα να συγχέονται με τις "48h" (16 δείγματα) και "96h" (9 δείγματα).

Συνολικά, το MLP αποδίδει πολύ καλά στις κατηγορίες με έντονες φασματικές διαφοροποιήσεις, όπως οι "96h" και "0h". Ωστόσο, στις ενδιάμεσες κατηγορίες, όπως οι "8h", "12h" και "24h", παρατηρούνται συγχύσεις λόγω μικρών διαφορών στα φασματικά χαρακτηριστικά. Το σημαντικό που πρέπει να σημειωθεί είναι ότι οι περισσότερες λανθασμένες ταξινομήσεις βρίσκονται σε γειτονικές κλάσεις, κάτι που δείχνει και τη στιβαρότητα και την ικανότητα γενίκευσης που έχει το συγκεκριμένο μοντέλο, έναντι των υπολοίπων.



Εικόνα 1.5.1-29. Μετρικές ταξινόμησης για κάθε μία από τις κλάσεις που προέκυψαν από το πείραμα με τα φυτά της τομάτας για τον αλγόριθμο MLP.

Ο πίνακας σύγκρισης (Εικόνα 1.5.1-28) και το διάγραμμα των μετρικών Precision, Recall και F1-Score (Εικόνα 1.5.1-29) αποτυπώνουν την απόδοση του μοντέλου MLP στην ταξινόμηση των φασματικών αποκρίσεων των φυτών τομάτας ανά χρονικό διάστημα. Από τον πίνακα σύγκρισης, παρατηρούμε ότι η κατηγορία "96h" ξεχωρίζει για την υψηλή της ακρίβεια, με 380 σωστές προβλέψεις και ελάχιστες λανθασμένες, κυρίως προς την κατηγορία "72h". Αυτή η επίδοση επιβεβαιώνεται από το διάγραμμα, όπου οι τιμές Precision, Recall και F1-Score για την "96h" είναι από τις υψηλότερες, προσεγγίζοντας το 1, κάτι που αναδεικνύει την έντονη διαφοροποίηση των χαρακτηριστικών στη φάση

αυτή.

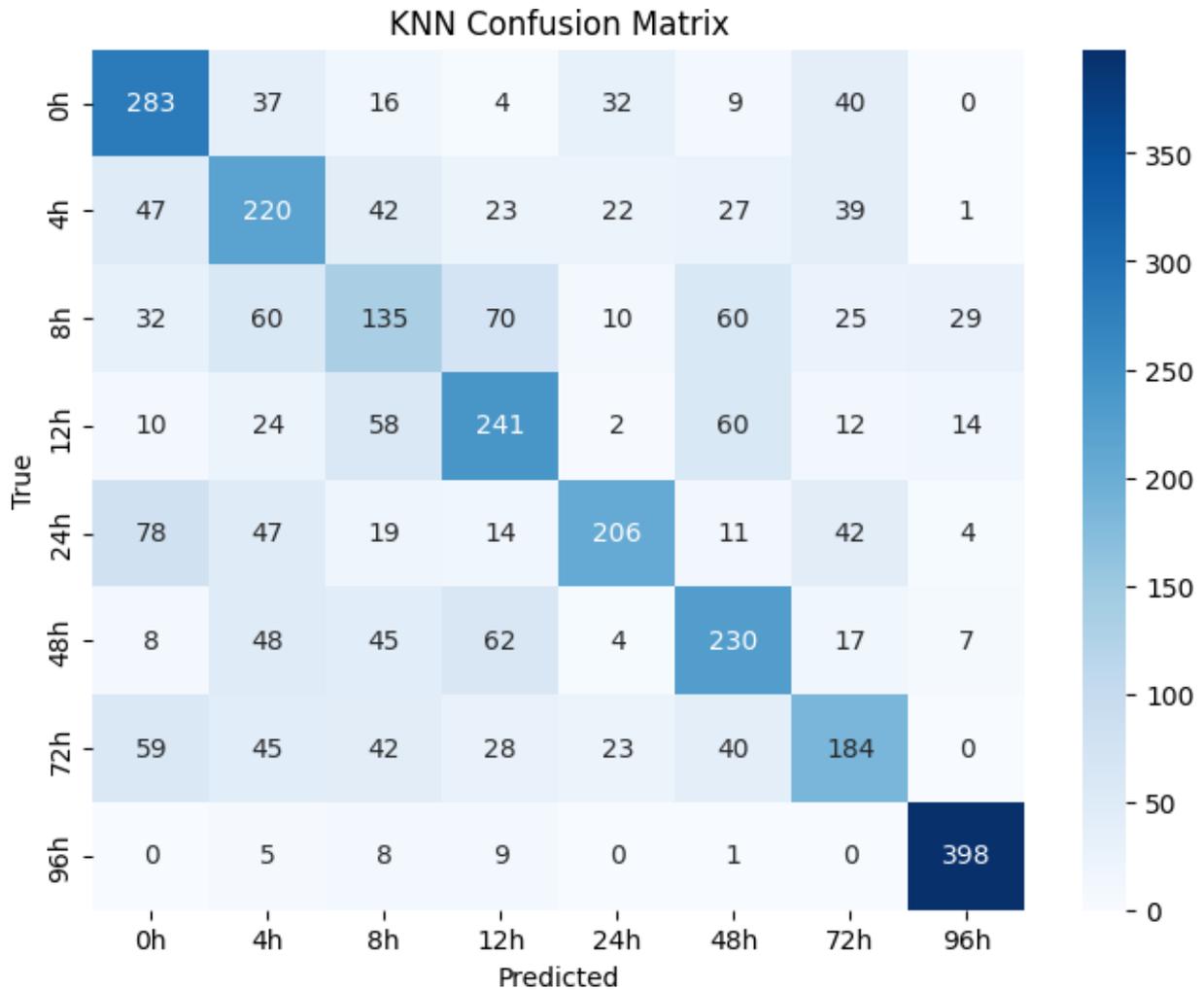
Ο πίνακας σύγκρισης (Εικόνα 1.5.1-30) και το διάγραμμα μετρικών (Εικόνα 1.5.1-31) αξιολογούν την απόδοση του αλγορίθμου k-Nearest Neighbors (kNN) για την κατηγοριοποίηση των φασματικών αποκρίσεων φυτών τομάτας σε επτά χρονικές κατηγορίες μετά τη μόλυνση (0h, 4h, 8h, 12h, 24h, 48h, 96h). Συνολικά, ο αλγόριθμος k-Nearest Neighbors (kNN) επιτυγχάνει υψηλή ακρίβεια στην κατηγορία "96h", με 398 σωστές προβλέψεις, λόγω της σαφούς φασματικής διαφοροποίησης αυτής της κατηγορίας σε σχέση με τις υπόλοιπες. Αντίθετα, οι κατηγορίες "8h" και "12h" παρουσιάζουν τις μεγαλύτερες προκλήσεις, λόγω αλληλεπικαλυπτόμενων χαρακτηριστικών.

Η κατηγορία "0h" εμφανίζει επίσης καλή απόδοση, με 283 σωστές προβλέψεις. Ωστόσο, παρατηρούνται αρκετές αναβαθμισμένες ταξινομήσεις προς την κατηγορία "4h". Παρόμοια, η κατηγορία "4h" παρουσιάζει σχετική σταθερότητα, αλλά ορισμένα δείγματα ταξινομούνται εσφαλμένα στις γειτονικές κατηγορίες "0h" και "8h".

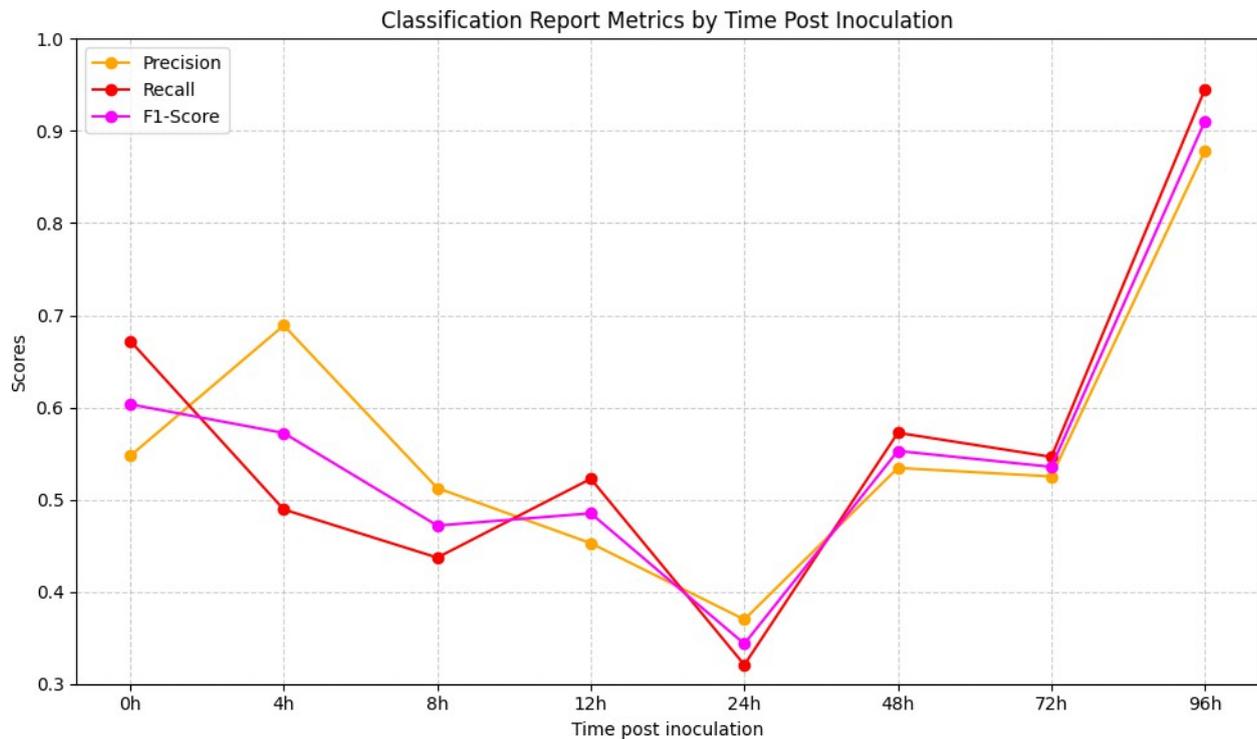
Η κατηγορία "24h" επιτυγχάνει 206 σωστές προβλέψεις, αλλά παρατηρούνται αρκετές λανθασμένες ταξινομήσεις στις κατηγορίες "12h" και "48h". Αντίστοιχα, η κατηγορία "48h" καταγράφει 230 σωστές προβλέψεις, ενώ η ταξινόμηση των κατηγοριών "48h" και "72h" επηρεάζεται από τη σχετική φασματική εγγύτητα, οδηγώντας σε μειωμένη ακρίβεια.

Στο διάγραμμα των στατιστικών μετρικών (Metrics), η κατηγορία "96h" διακρίνεται για την υψηλότερη απόδοση σε όλες τις μετρικές (Precision, Recall, F1-Score), ενώ οι κατηγορίες "8h" και "12h" εμφανίζουν τη χαμηλότερη απόδοση, υποδεικνύοντας προκλήσεις στην ακριβή διάκριση αυτών των χρονικών φάσεων. Οι κατηγορίες "24h" και "48h" παρουσιάζουν σταθερή βελτίωση, γεγονός που δείχνει ότι οι φασματικές διαφορές γίνονται πιο διακριτές με την πάροδο του χρόνου.

Συνολικά, τα αποτελέσματα αναδεικνύουν την αποτελεσματικότητα του kNN στις κατηγορίες με έντονες φασματικές διαφοροποιήσεις (π.χ. "96h"), αλλά και την ανάγκη βελτίωσης της απόδοσης σε πρώιμες χρονικές κατηγορίες, όπου τα φασματικά χαρακτηριστικά εμφανίζονται λιγότερο διακριτά. Προτείνεται η ενσωμάτωση πιο σύνθετων χαρακτηριστικών ή η χρήση άλλων μοντέλων για την ενίσχυση της ακρίβειας στις κατηγορίες με αλληλεπικαλυπτόμενα δεδομένα.



Εικόνα 1.5.1-30. Πίνακας σύγχυσης των αποτελεσμάτων των φασματικών αποκρίσεων των φυτών της τομάτας με τη μέθοδο KNN.



Εικόνα 1.5.1-31. Μετρικές ταξινόμησης για κάθε μία από τις κλάσεις που προέκυψαν από το πείραμα με τα φυτά της τομάτας για τον αλγόριθμο KNN.

Ο πίνακας σύγχυσης (Εικόνα 1.5.1-30) και το διάγραμμα μετρικών (Εικόνα 1.5.1-31) αξιολογούν την απόδοση του αλγορίθμου k-Nearest Neighbors (kNN) για την κατηγοριοποίηση των φασματικών αποκρίσεων φυτών τομάτας σε επτά χρονικές κατηγορίες μετά τη μόλυνση (0h, 4h, 8h, 12h, 24h, 48h, 96h). Συνολικά, ο αλγόριθμος επιτυγχάνει υψηλή ακρίβεια στην κατηγορία "96h", με 398 σωστές προβλέψεις, λόγω της σαφούς φασματικής διαφοροποίησης αυτής της κατηγορίας σε σχέση με τις υπόλοιπες. Αντίθετα, οι κατηγορίες "8h" και "12h" παρουσιάζουν τις μεγαλύτερες προκλήσεις, λόγω αλληλεπικαλυπτόμενων χαρακτηριστικών.

Η κατηγορία "0h" εμφανίζει επίσης καλή απόδοση, με 283 σωστές προβλέψεις. Ωστόσο, παρατηρούνται αρκετές αναβαθμισμένες ταξινομήσεις προς την κατηγορία "4h". Παρόμοια, η κατηγορία "4h" παρουσιάζει σχετική σταθερότητα, αλλά ορισμένα δείγματα ταξινομούνται εσφαλμένα στις γειτονικές κατηγορίες "0h" και "8h".

Η κατηγορία "24h" επιτυγχάνει 206 σωστές προβλέψεις, αλλά παρατηρούνται αρκετές λανθασμένες ταξινομήσεις στις κατηγορίες "12h" και "48h". Αντίστοιχα, η κατηγορία "48h" καταγράφει 230 σωστές προβλέψεις, ενώ η ταξινόμηση των κατηγοριών "48h" και "72h" επηρεάζεται από τη σχετική φασματική εγγύτητα, οδηγώντας σε μειωμένη ακρίβεια.

Στο διάγραμμα μετρικών, η κατηγορία "96h" διακρίνεται για την υψηλότερη απόδοση σε όλες τις μετρικές (Precision, Recall, F1-Score), ενώ οι κατηγορίες "8h" και "12h" εμφανίζουν τη χαμηλότερη απόδοση, υποδεικνύοντας προκλήσεις στην ακριβή διάκριση αυτών των χρονικών φάσεων. Οι κατηγορίες "24h" και "48h" παρουσιάζουν σταθερή βελτίωση, γεγονός που δείχνει ότι οι φασματικές διαφορές γίνονται πιο διακριτές με την

πάροδο του χρόνου.

Συνολικά, τα αποτελέσματα αναδεικνύουν την αποτελεσματικότητα του kNN στις κατηγορίες με έντονες φασματικές διαφοροποιήσεις (π.χ. "96h"), αλλά και την ανάγκη βελτίωσης της απόδοσης σε πρώιμες χρονικές κατηγορίες, όπου τα φασματικά χαρακτηριστικά εμφανίζονται λιγότερο διακριτά. Προτείνεται η ενσωμάτωση πιο σύνθετων χαρακτηριστικών ή η χρήση άλλων μοντέλων για την ενίσχυση της ακρίβειας στις κατηγορίες με αλληλεπικαλυπτόμενα δεδομένα.

3 ΣΥΝΟΨΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ

Η μελέτη εξετάζει τη χρήση υπερφασματικής απεικόνισης σε συνδυασμό με τεχνικές μηχανικής μάθησης (ML) για τη διάγνωση ασθενειών στα φυτά, με έμφαση στο ωίδιο της αμπέλου και τη τεφρά σήψη της τομάτας. Αξιολογούνται τρία μοντέλα ML: Random Forest - RF, Ο αλγόριθμος K-Nearest Neighbors (KNN) και το Πολυστρωματικό Perceptron (Multilayer Perceptron - MLP), βάσει μετρικών όπως η ακρίβεια, το precision, και το F1-score. Η μεθοδολογία περιλαμβάνει τη συλλογή υπερφασματικών δεδομένων από υγιή και μολυσμένα φυτά, σε συνθήκες εργαστηρίου και σε συνθήκες αγρού, σε προσυμπτωματικό στάδιο της μόλυνσης. Ακολούθησε προεπεξεργασία των δεδομένων, με αφαίρεση των ακραίων τιμών και τη δημιουργία ενός "pipeline", με διαφορετικές διαδοσμένες μεθόδους προεπεξεργασίας φασματικών δεδομένων. Ακολουθεί εξαγωγή χαρακτηριστικών (φασματικοί δείκτες) και η επιλογή των καταλληλότερων χαρακτηριστικών που θα χρησιμοποιηθούν κατά την εφαρμογή αλγορίθμων μηχανικής μάθησης για την ταξινόμηση, με τον αλγόριθμο Random Forest.

Για την άμπελο, τα δεδομένα ήταν σχετικά απλά, καθώς υπήρχαν μόνο 3 φασματικές κατηγορίες που σχετίζονταν με τις διαφορετικές καταστάσεις των φυτών (υγιή, ελαφρώς μολυσμένα, σοβαρά μολυσμένα). Αυτό είχε ως αποτέλεσμα και τα τρία μοντέλα να επιτυγχάνουν σχεδόν τέλεια αποτελέσματα, καταγράφοντας ακρίβεια που έφτανε το 99% τόσο στη φάση της εκπαίδευσης όσο και στη φάση των δοκιμών. Η απλότητα των δεδομένων επέτρεψε στα μοντέλα να αναγνωρίζουν με ακρίβεια τις φασματικές διαφοροποιήσεις, ενώ η συνολική απόδοση ήταν σχεδόν ισοδύναμη για όλα τα μοντέλα.

Στην περίπτωση της τομάτας, η κατάσταση ήταν πιο σύνθετη, καθώς τα δεδομένα περιλάμβαναν 8 φασματικές κατηγορίες που αντιπροσώπευαν το χρόνο μετά την τεχνητή μόλυνση. Η αυξημένη φασματική πολυπλοκότητα επηρέασε την απόδοση των μοντέλων. Το MLP αναδείχθηκε ως το καλύτερο μοντέλο, με την καλύτερη ικανότητα γενίκευσης. Η πλειοψηφία των λανθασμένων ταξινομήσεων αφορούσε γειτονικές φασματικές κατηγορίες, γεγονός που υποδηλώνει τη σταθερότητα του μοντέλου MLP και την ικανότητά του να διαχειρίζεται πιο περίπλοκα δεδομένα. Το μοντέλο RF είχε επίσης ικανοποιητική απόδοση, αλλά εμφάνισε μικρότερη ικανότητα γενίκευσης στις ενδιάμεσες κατηγορίες. Το μοντέλο KNN ήταν το μοντέλο με τη χαμηλότερη απόδοση, καθώς φάνηκε να δυσκολεύεται να διαχωρίσει φασματικά κοντινές κατηγορίες λόγω της αυξημένης πολυπλοκότητας.

4 ΒΙΒΛΙΟΓΡΑΦΙΑ

- Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, 13(1–2), 95–120.
- Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, 13(1–2), 95–120.
- Baranoski, G. V. G., & Rokne, J. G. (2001). Vc2001a, 491–505.
- Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., ... Thompson, T. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data.
- Behmann, J., Mahlein, A. K., Rumpf, T., Römer, C., & Plümer, L. (2015). A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. *Precision Agriculture*, 16(3), 239–260.
- Blackburn, G. A. (1998). Quantifying chlorophylls and carotenoids at leaf and canopy scales: An evaluation of some hyperspectral approaches. *Remote Sensing of Environment*, 66(3), 273–285.
- Blackburn, G. A. (1998). Quantifying chlorophylls and carotenoids at leaf and canopy scales: An evaluation of some hyperspectral approaches. *Remote Sensing of Environment*, 66(3), 273–285.
- Calonnec, A., Cartolaro, P., & Naulin, J. (2006). A host-pathogen simulation model: powdery mildew of grapevine. *Plant Pathology*, 55(4), 493–504.
- Calonnec, A., Cartolaro, P., Poupot, C., Dubourdieu, D., & Darriet, P. (2004). Effects of *Uncinula necator* on the yield and quality of grapes and wine. *Plant Pathology*, 53(4), 434–445.
- Chung, C.-L., Huang, K.-J., Chen, S.-Y., Lai, M.-H., Chen, Y.-C., & Kuo, Y.-F. (2016). Detecting *Bakanae* disease in rice seedlings by machine vision. *Computers and Electronics in Agriculture*, 121, 404–411.
- Datt, B. (1998). Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+ b, and total carotenoid content in eucalyptus leaves. *Remote Sensing of Environment*, 66(2), 111–121.
- Dawson, T. P., & Curran, P. J. (1998). Technical note A new technique for interpolating the reflectance red edge position.
- Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., ... & Foster, G. D. (2012). The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology*, 13(4), 414–430.
- Ebrahimi, M. A., Khoshtaghaza, M. H., Minaei, S., & Jamshidi, B. (2017). Vision-based pest detection based on SVM classification method. *Computers and Electronics in Agriculture*, 137, 52–58.
- Elmer, P. A., & Michailides, T. J. (2007). Epidemiology of *Botrytis cinerea* in orchard and vine crops. In *Botrytis: biology, pathology and control* (pp. 243–272). Dordrecht: Springer Netherlands.
- Feng, Y., Peng, Y., Cui, N., Gong, D., & Zhang, K. (2017). Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Computers and Electronics in Agriculture*, 136, 71–78.
- Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145(September 2017), 311–318.
- Fillinger, S., & Elad, Y. (Eds.). (2016). *Botrytis – the fungus, the pathogen and its management in agricultural systems*. Springer International Publishing.
- Gadoury, D. M., & Pearson, R. C. (1990). Germination of ascospores and infection of *Vitis* by *Uncinula necator*. *Phytopathology*, 80(11), 1198–1203.
- Gadoury, D. M., Seem, R. C., Ficke, A., & Wilcox, W. F. (2012). Grapevine powdery

- mildew (*Erysiphe necator*): a fascinating system for the study of the biology, ecology, and epidemiology of an obligate biotroph. *Molecular Plant Pathology*, 13(1), 1–16.
- Gaforio, L., Gomez, I., & Ortega, J. (2019). Biological control of grapevine powdery mildew using beneficial microorganisms. *Biological Control Journal*, 85(2), 112–120.
- Gitelson, A. A., & Merzlyak, M. N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing*, 18(12), 2691–2697.
- Golhani, K., Balasundram, S. K., Vadamalai, G., & Pradhan, B. (2018). A review of neural networks in plant disease detection using hyperspectral data. *Information Processing in Agriculture*, 5(3), 354–371.
- Gubler, W. D., Hirschfeld, D. J., & Thomas, C. S. (1996). Control of powdery mildew on grapes using sulfur and synthetic fungicides. *American Journal of Enology and Viticulture*, 47(2), 205–212.
- Herrmann, R., & Onkelinx, C. (1986). Quantities and units in clinical chemistry: Nebulizer and flame properties in flame emission and absorption spectrometry (Recommendations 1986). *Pure and Applied Chemistry*, 58(12), 1737–1742.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). 2002 Huete overview of the.pdf, 83, 195–213.
- Jacquemoud, S., & Ustin, S. L. (2001). Leaf optical properties: A state of the art. *8th International Symposium of Physical Measurements & Signatures in Remote Sensing*, 223–332.
- Jarvis, W. R. (1977). *Botryotinia and Botrytis species: Taxonomy, physiology, and pathogenicity*. Monograph No. 15. Canada Department of Agriculture.
- Jarvis, W. R. (1994). Latent infections in the pre- and postharvest environment. *Postharvest News and Information*, 5(3), 31N–34N.
- Jensen, J. R. (2007). *Remote sensing of the environment: an earth resource perspective*: Pearson Prentice Hall. *Upper Saddle River, NJ*.
- Karimi, Y., Prasher, S. O., Patel, R. M., & Kim, S. H. (2006). Application of support vector machine technology for weed and nitrogen stress detection in corn. *Computers and Electronics in Agriculture*, 51(1–2), 99–109.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors (Switzerland)*, 18(8), 1–29.
- LÓPEZ-GRANADOS, F., PEÑA-BARRAGÁN, J. M., JURADO-EXÓSITO, M., FRANCISCO-FERNÁNDEZ, M., Cao, R., ALONSO-BETANZOS, A., & FONTENLA-ROMERO, O. (2008). Multispectral classification of grass weeds and wheat (*Triticum durum*) using linear and nonparametric functional discriminant analysis and neural networks. *Weed Research*, 48(1), 28–37.
- Martens, H., & Næs, T. (1989). *Multivariate Calibration*. John Wiley & Sons.
- Mishra, P., Asaari, M. S. M., Herrero-Langreo, A., Lohumi, S., Diezma, B., & Scheunders, P. (2017). Close range hyperspectral imaging of plants: A review. *Biosystems Engineering*, 164, 49–67.
- Morellos, A., Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., ... Mouazen, A. M. (2016). Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosystems Engineering*, 152, 104–116.
- Moshou, D., Bravo, C., Oberti, R., West, J. S., Ramon, H., Vougioukas, S., & Bochtis, D. (2011). Intelligent multi-sensor system for the detection and treatment of fungal diseases in arable crops. *Biosystems Engineering*, 108(4), 311–321.
- Moshou, D., Bravo, C., Oberti, R., West, J., Bodria, L., McCartney, A., & Ramon, H. (2005). Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. *Real-Time Imaging*, 11(2),

75–83.

- Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A., & Ramon, H. (2004). Automatic detection of “yellow rust” in wheat using reflectance measurements and neural networks. *Computers and Electronics in Agriculture*, 44(3), 173–188.
- Moshou, D., Pantazi, X.-E., Kateris, D., & Gravalos, I. (2014a). Water stress detection based on optical multisensor fusion with a least squares support vector machine classifier. *Biosystems Engineering*, 117, 15–22.
- Nahvi, B., Habibi, J., Mohammadi, K., Shamshirband, S., & Al Razgan, O. S. (2016). Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature. *Computers and Electronics in Agriculture*, 124, 150–160.
- Oni, F. E., Esmaeel, Q., Onyeka, J. T., Adeleke, R., Jacquard, C., Clement, C., ... & Höfte, M. (2022). Pseudomonas lipopeptide-mediated biocontrol: chemotaxonomy and biological activity. *Molecules*, 27(2), 372.
- Pantazi, X. E., Moshou, D., Oberti, R., West, J., Mouazen, A. M., & Bochtis, D. (2017). Detection of biotic and abiotic stresses in crops by using hierarchical self organizing classifiers. *Precision Agriculture*, 18(3), 383–393.
- Pantazi, X. E., Tamouridou, A. A., Alexandridis, T. K., Lagopodi, A. L., Kontouris, G., & Moshou, D. (2017). Detection of Silybum marianum infection with Microbotryum silybum using VNIR field spectroscopy. *Computers and Electronics in Agriculture*, 137, 130–137.
- Pappas, A. C. (2000). Epidemiology and control of Botrytis cinerea in tomato crops grown under cover in Greece. *EPPO Bulletin*, 30(2), 269–274.
- Patil, A. P., & Deka, P. C. (2016). An extreme learning machine approach for modeling evapotranspiration using extrinsic inputs. *Computers and Electronics in Agriculture*, 121, 385–392.
- Pearson, R. C., & Gadoury, D. M. (1987). Cleistothecia, the source of primary inoculum for grape powdery mildew in New York. *Phytopathology*, 77(11), 1509–1514.
- Pimentel, D., Berner, M., & Garcia, P. (2020). Early detection of grapevine powdery mildew using hyperspectral imaging and machine learning algorithms. *Precision Agriculture*, 21(3), 562–578.
- Raza, S. E. A., Prince, G., Clarkson, J. P., & Rajpoot, N. M. (2020). Automatic detection of diseased tomato plants using thermal and stereo visible light images. *Scientific Reports*, 10(1), 1–11.
- Rinnan, Å., van den Berg, F., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28(10), 1201–1222.
- Rouse, J. W., Hass, R. H., Schell, J. A., & Deering, D. W. (1972). Monitoring Vegetation Systems in the Great Plains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*.
- Sakudo, A., Sukanuma, Y., Kobayashi, T., Onodera, T., & Ikuta, K. (2006). Near-infrared spectroscopy: Promising diagnostic tool for viral infections. *Biochemical and Biophysical Research Communications*, 341(2), 279–284.
- Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8), 1627–1639.
- Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, 81(2), 337–354.
- Strange, R. N., & Scott, P. R. (2005). Plant Disease: A Threat to Global Food Security. *Annual Review of Phytopathology*, 43(1), 83–116.
- Suzuki, Y., Okamoto, H., & Kataoka, T. (2008). Image segmentation between crop and weed using hyperspectral imaging for weed detection in soybean field.

Environmental Control in Biology, 46(3), 163–173.

- Tamouridou, A. A., Alexandridis, T. K., Pantazi, X. E., Lagopodi, A. L., Kashefi, J., & Moshou, D. (2017). Evaluation of UAV imagery for mapping *Silybum marianum* weed patches. *International Journal of Remote Sensing*, 38(8–10), 2246–2259.
- Thanarajoo, S. S., Kong, L. L., Kadir, J., Lau, W. H., & Vadamalai, G. (2014). Detection of Coconut cadang-cadang viroid (CCCVd) in oil palm by reverse transcription loop-mediated isothermal amplification (RT-LAMP). *Journal of Virological Methods*, 202, 19–23.
- Tzitzikas, A., Papadopoulou, A., & Katsantonis, D. (2008). Geographical distribution and epidemiology of grapevine powdery mildew in Greece. *Journal of Plant Pathology*, 90(1), 45–52.
- Vogelmann, J. E., Rock, B. N., & Moss, D. M. (1993). Red edge spectral measurements from sugar maple leaves. *Remote Sensing*, 14(8), 1563–1575.
- Wang, X., Zhang, M., Zhu, J., & Geng, S. (2008). Spectral prediction of *Phytophthora infestans* infection on tomatoes using artificial neural network (ANN). *International Journal of Remote Sensing*, 29(6), 1693–1706.
- Wilcox, W. F., Gadoury, D. M., & Seem, R. C. (2015). Grapevine powdery mildew: Variability in fungicide sensitivity, and strategies for disease management. *Plant Disease*, 99(10), 1287–1297.
- Williamson, B., Tudzynski, B., Tudzynski, P., & Van Kan, J. A. L. (2007). *Botrytis cinerea*: the cause of grey mould disease. *Molecular Plant Pathology*, 8(5), 561–580.