



**ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΕΛΟΠΟΝΝΗΣΟΥ**
University of the Peloponnese

ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ
Π.Μ.Σ. ΣΤΗΝ ΕΠΙΣΤΗΜΗ ΥΠΟΛΟΓΙΣΤΩΝ

Διπλωματική Εργασία

**Ενσωμάτωση Μηχανικής Μάθησης σε Γεωγραφικά Συστήματα
Πληροφοριών για την Ανάλυση και Πρόβλεψη Γεωχωρικών
Δεδομένων**

Ανδρέου Δημήτριος

A.M.: 2022202302001

Επιβλέπων: Βασιλάκης Κωνσταντίνος

Τρίπολη | 2025

Περιεχόμενα

Ευρετήριο Εικόνων	4
Ευρετήριο Πινάκων.....	5
Περίληψη	6
Abstract.....	7
1 Εισαγωγή.....	8
2 Γεωχωρικά Δεδομένα	9
2.1 Εισαγωγή	9
2.2 Διάκριση γεωχωρικών δεδομένων.....	9
2.2.1 Περιγραφικά Δεδομένα	9
2.2.2 Γεωγραφικά Δεδομένα	9
3 Γεωγραφικά Συστήματα Πληροφοριών.....	11
3.1 Εισαγωγή	11
3.2 Στάδια και διαδικασίες στα ΓΣΠ.....	11
3.2.1 Καθορισμός του προβλήματος	12
3.2.2 Διαδικασία μετατροπής των στοιχείων σε πληροφορία	12
3.2.3 Διαδικασία εξαγωγής συμπερασμάτων	13
4 Μηχανική Μάθηση	15
4.1 Είδη Μηχανικής Μάθησης	15
4.1.1 Επιβλεπόμενη Μηχανική Μάθηση	15
4.1.2 Μη επιβλεπόμενη Μηχανική Μάθηση.....	20
4.1.3 Ενισχυτική Μάθηση	21
4.2 Νευρωνικά Δίκτυα.....	22
5 Τηλεπισκόπηση.....	24
5.1 Πρόγραμμα Copernicus.....	24
5.2 Sentinel 2.....	24
6 Συνδυασμός GIS και Μηχανικής μάθησης	27
7 Περιπτώσεις χρήσης	28
7.1 Παγκόσμια Προγράμματα Χαρτογράφησης Καλύψεων και Χρήσεων Γης.....	28
7.2 Γεωργία Ακριβείας και Μηχανική Μάθηση	29
7.3 Επιχειρηματική Νοημοσύνη μέσω Τηλεπισκόπησης και Μηχανικής Μάθησης	29

7.4	Ανίχνευση Οροφών για Φωτοβολταϊκά Συστήματα μέσω Μηχανικής Μάθησης.....	30
8	Μελέτη περίπτωσης	32
8.1	Γεωγραφικά στοιχεία	32
8.1.1	Τοπογραφικά Στοιχεία	32
8.1.2	Δημογραφικά Στοιχεία.....	33
8.1.3	Υδρολογικά Στοιχεία-Κλίμα.....	33
8.1.4	Οικονομικές Δραστηριότητες	34
8.2	Χρήσεις Γης.....	34
8.3	Λήψη εικόνων.....	35
8.4	Επαναδειγματοληψία και αποκοπή εικόνας	36
8.5	Εργαλεία και Βιβλιοθήκες	40
8.5.1	Python	40
8.5.2	Numpy.....	40
8.5.3	Scikit-learn.....	40
8.5.4	Geospatial Data Abstraction Library	41
8.5.5	Semi-Automatic Classification Plugin (SCP)	41
8.6	Επιλογή Δεδομένων Εκπαίδευσης	42
8.7	Κατηγοριοποιήσεις.....	44
8.7.1	Κατηγοριοποίηση ελάχιστης απόστασης	46
8.7.2	Κατηγοριοποίηση Μέγιστης Πιθανοφάνειας	47
8.7.3	Κατηγοριοποίηση φασματικής γωνίας.....	48
8.8	Αξιολόγηση – Πολύγωνα Ελέγχου.....	50
8.8.1	Πίνακας Σύγχυσης Ελάχιστης Απόστασης	50
8.8.2	Πίνακας Σύγχυσης Μέγιστης Πιθανοφάνειας	51
8.8.3	Πίνακας Σύγχυσης Φασματικής Γωνίας.....	53
9	Συμπεράσματα.....	55
10	Επίλογος.....	56
11	Βιβλιογραφία	57

Ευρετήριο Εικόνων

Εικόνα 1: Γεωχωρικά δεδομένα.....	10
Εικόνα 2: Στάδια και διαδικασίες ενός Γ.Σ.Π. (Κουτσόπουλος 2002).....	12
Εικόνα 3: Είδη και Εφαρμογές της Μηχανικής Μάθησης	15
Εικόνα 4: Κατηγοριοποίηση - Παλινδρόμηση	16
Εικόνα 5: Ταξινόμηση Μέγιστης Πιθανοφάνειας.....	17
Εικόνα 6: Ταξινόμηση Φασματικής Γωνίας	18
Εικόνα 7: Ταξινόμηση Ελάχιστης Απόστασης.....	18
Εικόνα 8: Γραμμική Παλινδρόμηση	19
Εικόνα 9: Δέντρα Αποφάσεων για Παλινδρόμηση	20
Εικόνα 10: Random Forest.....	20
Εικόνα 11. Μη επιβλεπόμενη μάθηση	21
Εικόνα 12: Ενισχυτική Μάθηση	22
Εικόνα 13: Είδη Μηχανικής Μάθησης.....	22
Εικόνα 14: Νευρωνικό Δίκτυο.....	23
Εικόνα 15: Γεωγραφική θέση του νομού Κορινθίας (πηγή elwikipedia.org)	32
Εικόνα 16: Γεωφυσική απεικόνιση του νομού Κορινθίας (πηγή Google Maps)	33
Εικόνα 17: Χαρτοσύνθεση περιοχής μελέτης με όμορους Δήμους	34
Εικόνα 18: Χαρτοσύνθεση με δορυφορική εικόνα και χρήσεις γης CLC 2018.....	35
Εικόνα 19: Επιλογή πολυγώνου επιθυμητής περιοχής και δορυφόρου	36
Εικόνα 20: Διαθέσιμες δορυφορικές φωτογραφίες.....	36
Εικόνα 21: Αναδόμηση στο κανάλι B2.....	37
Εικόνα 22: Αρχική Δορυφορική Εικόνα	38
Εικόνα 23: Η αποκοπή των εικόνων	39
Εικόνα 24: Κανάλια που παρέμειναν	40
Εικόνα 25: Δημιουργία Πολυγώνων	42
Εικόνα 26: Πολύγωνα που δημιουργήθηκαν	42
Εικόνα 27: Στατιστικά στοιχεία δεδομένων εκπαίδευσης	43
Εικόνα 28: Επιλογή Classification στο SCP	45
Εικόνα 29: Επιλογή Αλγορίθμου	45
Εικόνα 30: Ταξινόμηση ελάχιστης απόστασης.....	46
Εικόνα 31: Ταξινόμηση μέγιστης πιθανοφάνειας	47
Εικόνα 32: Σημαντικά σφάλματα (με διαφορετικό χρώμα κάθε περίπτωση)	48
Εικόνα 33. Κατηγοριοποίηση φασματικής γωνίας.....	48
Εικόνα 34: Σημαντικά σφάλματα (με διαφορετικό χρώμα κάθε περίπτωση)	49

Ευρετήριο Πινάκων

Πίνακας 1: Κανάλια Sentinel 2.....	26
Πίνακας 2: Κατηγορίες χρήσεων γης CLC 2018 που υπάρχουν στην περιοχή μελέτης.	35
Πίνακας 3: Πίνακας σύγκυσης ελάχιστης απόστασης.....	50
Πίνακας 4: Πίνακας σύγκυσης μέγιστης πιθανοφάνειας	52
Πίνακας 5: Πίνακας σύγκυσης φασματικής γωνίας	53

Περίληψη

Η παρούσα διπλωματική εργασία διερευνά τη συνδυαστική εφαρμογή τεχνικών μηχανικής μάθησης και γεωγραφικών συστημάτων πληροφοριών (GIS) για την ανάλυση και πρόβλεψη γεωχωρικών δεδομένων, με στόχο την ενίσχυση της αποτελεσματικότητας και της ακρίβειας στην επεξεργασία τηλεπισκοπικών πληροφοριών. Στην εποχή της πληθωρικής γεωχωρικής πληροφορίας και της αυξανόμενης ανάγκης για άμεση λήψη τεκμηριωμένων αποφάσεων, η ανάπτυξη μεθοδολογιών που συνδυάζουν την επεξεργαστική ισχύ της μηχανικής μάθησης με τις δομημένες γεωγραφικές βάσεις των GIS αποτελεί πρόκληση και συνάμα ευκαιρία.

Η εργασία αναπτύσσει ένα θεωρητικό υπόβαθρο σχετικά με τα είδη γεωχωρικών δεδομένων, τη δομή και τις λειτουργίες των GIS, και τις βασικές αρχές της επιβλεπόμενης και μη επιβλεπόμενης μάθησης, εστιάζοντας ιδιαίτερα σε τεχνικές ταξινόμησης χρήσεων γης. Παράλληλα, αναδεικνύονται οι προκλήσεις που ενυπάρχουν στη χρήση τηλεπισκοπικών δεδομένων και εξετάζονται οι δυνατότητες και οι περιορισμοί της μηχανικής μάθησης όταν εφαρμόζεται σε γεωχωρικές αναλύσεις.

Για τη μελέτη περίπτωσης, χρησιμοποιήθηκαν δορυφορικά δεδομένα Sentinel-2 για την περιοχή της Κορινθίας. Μέσα από διαδικασίες επαναδειγματοληψίας, επιλογής κατάλληλων καναλιών, δημιουργίας δεδομένων εκπαίδευσης και εφαρμογής αλγορίθμων ταξινόμησης (Μέγιστης Πιθανοφάνειας, Φασματικής Γωνίας και Ελάχιστης Απόστασης), αξιολογήθηκαν τα αποτελέσματα τόσο ποσοτικά μέσω πινάκων σύγχυσης όσο και ποιοτικά μέσω οπτικής ανάλυσης.

Τα ευρήματα αποδεικνύουν ότι η χρήση τεχνικών μηχανικής μάθησης σε GIS βελτιώνει σημαντικά την ποιότητα και τη διακριτική ικανότητα της ταξινόμησης χρήσεων γης. Ωστόσο, η επιτυχία της διαδικασίας εξαρτάται σε μεγάλο βαθμό από τη σωστή επιλογή δεδομένων εκπαίδευσης, την ποιότητα της φασματικής πληροφορίας και την κατανόηση των ιδιοτήτων κάθε αλγορίθμου. Η εργασία καταλήγει στο συμπέρασμα ότι η ενσωμάτωση μηχανικής μάθησης σε γεωγραφικά συστήματα πληροφοριών δεν αποτελεί μόνο τεχνική καινοτομία αλλά καθίσταται αναγκαίο εργαλείο για τη μελλοντική εξέλιξη της γεωχωρικής επιστήμης και εφαρμοσμένης χαρτογράφησης.

Λέξεις-κλειδιά: Μηχανική Μάθηση, Γεωγραφικά Συστήματα Πληροφοριών, Τηλεπισκόπηση, Ταξινόμηση Εικόνων, GeoAI, Χρήσεις Γης, Sentinel-2

Abstract

This thesis explores the combined application of machine learning techniques and Geographic Information Systems (GIS) for the analysis and prediction of geospatial data, aiming to enhance the efficiency and accuracy of remote sensing data processing. In an era characterized by an overwhelming abundance of geospatial information and an increasing need for timely and well-informed decision-making, the development of methodologies that combine the computational power of machine learning with the structured geographical databases of GIS emerges both as a challenge and an opportunity.

The study develops a solid theoretical foundation regarding the types of geospatial data, the structure and functionalities of GIS, and the fundamental principles of supervised and unsupervised learning, focusing particularly on land use classification techniques. Moreover, it highlights the challenges inherent in using remote sensing data and examines the capabilities and limitations of machine learning when applied to geospatial analysis.

For the case study, Sentinel-2 satellite data from the Corinthia region were utilized. Through processes such as resampling, appropriate band selection, creation of training data, and the application of classification algorithms (Maximum Likelihood, Spectral Angle Mapping, and Minimum Distance), the results were evaluated both quantitatively, through confusion matrices, and qualitatively, through visual analysis.

The findings demonstrate that the integration of machine learning techniques into GIS significantly improves the quality and discriminative power of land use classifications. Nevertheless, the success of the process heavily depends on the careful selection of training data, the quality of spectral information, and a deep understanding of the particularities of each algorithm. The thesis concludes that integrating machine learning into geographic information systems is not merely a technical innovation but is becoming an indispensable tool for the future advancement of geospatial science and applied cartography.

Keywords: Machine Learning, Geographic Information Systems (GIS), Remote Sensing, Image Classification, GeoAI, Land Use, Sentinel-2

1 Εισαγωγή

Η σύγχρονη εποχή χαρακτηρίζεται από μια πρωτοφανή αύξηση του όγκου γεωχωρικών δεδομένων, αποτέλεσμα της εκτεταμένης χρήσης δορυφορικών αποστολών, αισθητήρων τηλεπισκόπησης, συστημάτων GPS και άλλων τεχνολογιών γεωχωρικής παρατήρησης. Σε αυτό το πλαίσιο, η διαχείριση, ανάλυση και αξιοποίηση των δεδομένων αυτών με σκοπό την εξαγωγή χρήσιμων πληροφοριών καθίσταται όλο και πιο κρίσιμη. Τα Γεωγραφικά Συστήματα Πληροφοριών (GIS), που παραδοσιακά αποτελούσαν βασικό μέσο οργάνωσης και ανάλυσης γεωχωρικών πληροφοριών, αντιμετωπίζουν πλέον την πρόκληση της διαχείρισης δεδομένων μεγάλης κλίμακας και υψηλής πολυπλοκότητας.

Παράλληλα, η ραγδαία εξέλιξη της μηχανικής μάθησης έχει δημιουργήσει νέα εργαλεία που επιτρέπουν την αναγνώριση προτύπων, τη δημιουργία προβλεπτικών μοντέλων και την αυτοματοποιημένη λήψη αποφάσεων. Οι τεχνικές μηχανικής μάθησης, όταν ενσωματώνονται σε περιβάλλοντα GIS, προσφέρουν τη δυνατότητα πιο αποτελεσματικής διαχείρισης μεγάλων δεδομένων, βελτίωσης της ακρίβειας ταξινομήσεων και πρόβλεψης σύνθετων γεωχωρικών φαινομένων.

Η παρούσα εργασία στοχεύει στην εις βάθος μελέτη της αλληλεπίδρασης μεταξύ μηχανικής μάθησης και GIS, τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο. Στόχος είναι να αναδειχθεί η προστιθέμενη αξία της ενσωμάτωσης των δύο τεχνολογιών στην ανάλυση γεωχωρικών δεδομένων, εξετάζοντας τόσο τις δυνατότητες όσο και τους περιορισμούς που προκύπτουν στην πράξη. Μέσα από μια ολοκληρωμένη επισκόπηση των θεωρητικών αρχών που διέπουν τα GIS, τα είδη γεωχωρικών δεδομένων, τις τεχνικές επιβλεπόμενης και μη επιβλεπόμενης μηχανικής μάθησης, και τις μεθόδους ταξινόμησης τηλεπισκοπικών εικόνων, διαμορφώνεται ένα ενιαίο πλαίσιο κατανόησης του πεδίου.

Η εφαρμογή πραγματοποιείται στην περιοχή της Κορινθίας, επιλέγοντας δεδομένα υψηλής ανάλυσης από τον δορυφόρο Sentinel-2 και χρησιμοποιώντας σύγχρονα εργαλεία ανάλυσης, όπως το Semi-Automatic Classification Plugin (SCP) και βιβλιοθήκες μηχανικής μάθησης της Python. Η εργασία αναπτύσσει μεθοδολογίες επεξεργασίας και ταξινόμησης γεωχωρικών δεδομένων, αξιολογεί τις επιδόσεις διαφόρων αλγορίθμων και συζητά τα αποτελέσματα σε σχέση με την ποιότητα των δεδομένων εκπαίδευσης και την πολυπλοκότητα του αντικειμένου.

Η ερευνητική αυτή προσέγγιση αντανακλά μια σύγχρονη τάση προς την ενσωμάτωση τεχνικών τεχνητής νοημοσύνης στη γεωπληροφορική επιστήμη και αποσκοπεί στην ανάδειξη νέων δρόμων για την ανάλυση, πρόβλεψη και διαχείριση του γεωγραφικού χώρου. Τα συμπεράσματα της εργασίας συμβάλλουν τόσο στην ακαδημαϊκή συζήτηση για τη μελλοντική κατεύθυνση της γεωπληροφορικής όσο και στη βελτίωση πρακτικών εφαρμογών σε τομείς όπως η διαχείριση φυσικών πόρων, η γεωργία ακριβείας και η πολεοδομία.

2 Γεωχωρικά Δεδομένα

2.1 Εισαγωγή

Ως γεωχωρικά δεδομένα ορίζονται οποιαδήποτε δεδομένα αφορούν άμεσα ή έμμεσα, συγκεκριμένη τοποθεσία ή γεωγραφική περιοχή. Τα γεωχωρικά δεδομένα είναι πληροφορίες που περιγράφουν την τοποθεσία, τη σχετική θέση και τα χαρακτηριστικά αντικειμένων και φαινομένων που βρίσκονται στην επιφάνεια της Γης. Αυτά τα δεδομένα μπορούν να περιλαμβάνουν γεωγραφικές τοποθεσίες (συντεταγμένες γεωγραφικού μήκους και πλάτους), αλλά και πολυδιάστατες πληροφορίες όπως υψόμετρο, χρονικές στιγμές, και άλλα χαρακτηριστικά σχετικά με φυσικά και ανθρωπογενή αντικείμενα και δραστηριότητες στην επιφάνεια της Γης. Η ανάλυση γεωχωρικών δεδομένων επιτρέπει την κατανόηση προτύπων, σχέσεων και τάσεων, παρέχοντας σημαντικές πληροφορίες για τη λήψη αποφάσεων και τη σχεδίαση. (Κάβουρας Μ. 2015)

2.2 Διάκριση γεωχωρικών δεδομένων.

Τα γεωχωρικά δεδομένα μπορούν να χωριστούν σε 2 μεγάλες κατηγορίες: Τα Περιγραφικά Δεδομένα (Attribute Data) και τα Γεωγραφικά Δεδομένα (Geospatial Data).

2.2.1 Περιγραφικά Δεδομένα

Τα περιγραφικά δεδομένα, αναφέρονται σε πληροφορίες που περιγράφουν τις ιδιότητες, τα χαρακτηριστικά ή την ποιότητα γεωχωρικών αντικειμένων. Αυτά τα δεδομένα δεν περιλαμβάνουν τις τοποθεσίες ή τις γεωμετρίες των αντικειμένων (όπως σημεία, γραμμές, και πολύγωνα), αλλά τις πληροφορίες που περιγράφουν τα αντικείμενα αυτά. Τα περιγραφικά δεδομένα επιτρέπουν την ανάλυση και την ερμηνεία των γεωχωρικών δεδομένων με πιο πλήρη και πολύπλοκο τρόπο, καθώς προσθέτουν πληροφοριακό πλούτο και βάθος στην απλή γεωμετρική αναπαράσταση των αντικειμένων στην επιφάνεια της Γης. Για παράδειγμα, σε μια βάση δεδομένων που αφορά το δίκτυο υδρογραφίας, τα περιγραφικά δεδομένα μπορεί να περιλαμβάνουν το όνομα του ποταμού, το μήκος, την ποιότητα του νερού, τον τύπο ροής, και άλλα σχετικά δεδομένα. (Κάβουρας Μ. 2015)

2.2.2 Γεωγραφικά Δεδομένα

Τα γεωγραφικά δεδομένα αφορούν την αναπαράσταση και την περιγραφή των σημείων ή των φαινομένων πάνω ή κοντά στην επιφάνειά της. Περιλαμβάνουν πληροφορίες σχετικά με τοποθεσίες (συντεταγμένες), σχήματα και τοπολογίες (δηλαδή, τη διάταξη και τη σχετική θέση των γεωγραφικών αντικειμένων) και μπορούν να περιγράψουν φυσικές οντότητες (π.χ., ποτάμια, λίμνες, βουνά) καθώς και ανθρωπογενείς οντότητες (π.χ., δρόμους, κτίρια, πόλεις). Όσον αφορά την αναπαράσταση των γεωγραφικών δεδομένων, αυτά μπορούν να χωριστούν σε δύο επιμέρους κατηγορίες, τα διανυσματικά δεδομένα (vector data) και τα ψηφιδωτά δεδομένα (raster data). (Κάβουρας Μ. 2015)

2.2.2.1 Διανυσματικά Δεδομένα

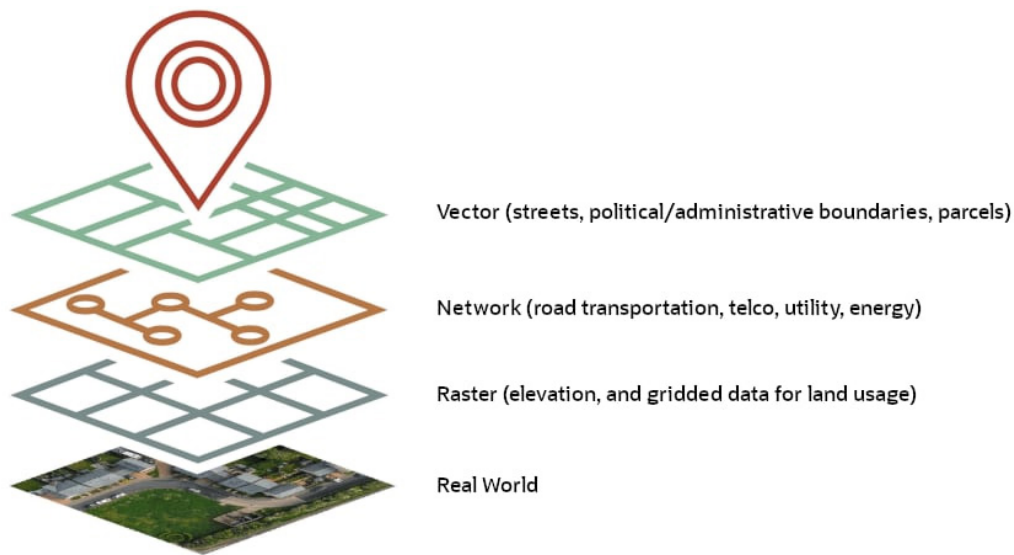
Τα διανυσματικά δεδομένα αναπαριστούν τον υπό μοντελοποίηση κόσμο μέσω σημείων, γραμμών, και πολυγώνων. Ενδεικτικά, τα σημεία μπορούν να αναπαραστήσουν σημεία ενδιαφέροντος (π.χ. αξιοθέατα, οικίες ή και μεγαλύτερους σχηματισμούς, π.χ. πόλεις), οι γραμμές μπορούν να

αντιπροσωπεύουν δρόμους και ποτάμια, ενώ τα πολύγωνα μπορούν να περιγράψουν λίμνες, δασικές εκτάσεις ή τα όρια μιας πόλης.

2.2.2.2 Ψηφιδωτά δεδομένα

Τα ψηφιδωτά δεδομένα αναπαριστούν τον υπό μοντελοποίηση κόσμο μέσω ενός κανονικού πλέγματος κελιών ή εικονοστοιχείων (pixel), με κάθε κελί να περιέχει μια τιμή που αντιπροσωπεύει ένα συγκεκριμένο χαρακτηριστικό ή μέτρηση της επιφάνειας της Γης, όπως υψόμετρο, θερμοκρασία, ή κάλυψη γης. Τα ψηφιδωτά δεδομένα χρησιμοποιούνται συχνά στην τηλεπισκόπηση και στις εικόνες από δορυφόρους.

Σε ένα γεωγραφικό σύστημα πληροφοριών, τυπικά χρησιμοποιούνται και τα δύο είδη δεδομένων τα οποία διαρθρώνονται σε επάλληλα επίπεδα, όπως φαίνεται στην Εικόνα 1.



Εικόνα 1: Γεωχωρικά δεδομένα

3 Γεωγραφικά Συστήματα Πληροφοριών

3.1 Εισαγωγή

Τα Γεωγραφικά Συστήματα Πληροφοριών (ΓΣΠ) (Geographic Information Systems / GIS) αποτελούν ένα τεχνολογικό εργαλείο που συλλέγει, αποθηκεύει, αναλύει, διαχειρίζεται και παρουσιάζει δεδομένα που σχετίζονται με τον χώρο και τη γεωγραφία. Διευκολύνουν την ερμηνεία, την ανάλυση, την οπτικοποίηση και την κατανόηση των σχέσεων, των μοτίβων και των τάσεων των γεωχωρικών δεδομένων, μέσω της χρήσης χαρτών και άλλων γραφικών αναπαραστάσεων. (Καπαγερίδης Ι., 2006).

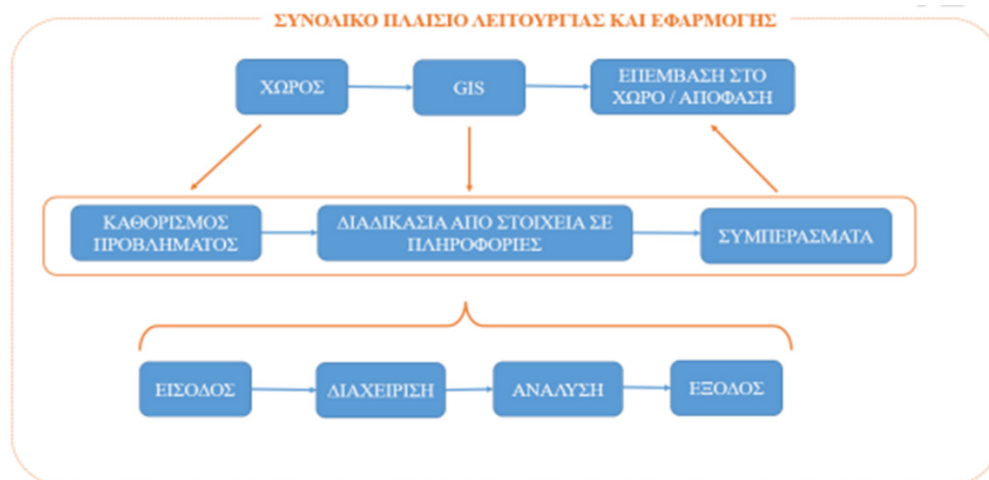
Τα ΓΣΠ εφαρμόζονται σε πληθώρα τομέων και πεδίων, όπως:

- Περιβαλλοντικές Επιστήμες: Στην παρακολούθηση και διαχείριση φυσικών πόρων, στην αξιολόγηση κινδύνων καταστροφών, στην προστασία του περιβάλλοντος.
- Πολεοδομία και Χωροταξία: Στον σχεδιασμό και ανάπτυξη των αστικών και αγροτικών χώρων, στη διαχείριση ακινήτων.
- Μεταφορές: Στην ανάλυση δικτύων μεταφοράς, στην βελτιστοποίηση διαδρομών, στον προγραμματισμό δημόσιων μεταφορών.
- Δημόσια Υγεία: Στην επιδημιολογία και την ανάλυση της διασποράς ασθενειών, στον προγραμματισμό υγειονομικών υπηρεσιών.
- Εκπαίδευση: Στην έρευνα και την ανάπτυξη εκπαιδευτικών προγραμμάτων που σχετίζονται με τη γεωγραφία και την περιβαλλοντική εκπαίδευση.
- Άμυνα και Εθνική Ασφάλεια: Για την ανάλυση του εδάφους, στρατηγικής σχεδίασης και επιχειρησιακής έρευνας.

3.2 Στάδια και διαδικασίες στα ΓΣΠ

Για την αποτελεσματική χρήση των ΓΣΠ και την εξαγωγή ορθών αποτελεσμάτων, έχει διαμορφωθεί ένα συνολικό πλαίσιο λειτουργίας και εφαρμογής τους, το οποίο περιλαμβάνει τις ακόλουθες διαδικασίες (Εικόνα 2):

- καθορισμός του προβλήματος,
- μετατροπή των στοιχείων σε πληροφορία και
- εξαγωγή συμπερασμάτων.



Εικόνα 2: Στάδια και διαδικασίες ενός Γ.Σ.Π. (Κουτσόπουλος 2002)

3.2.1 Καθορισμός του προβλήματος

Ο ορθός και πλήρης καθορισμός του προβλήματος είναι κρίσιμης σημασίας για την φάση για την επιτυχή εφαρμογή του ΓΣΠ. Ο στόχος της διαδικασίας αυτής είναι να καθοριστεί το πρόβλημα με σαφήνεια, ορθότητα και πληρότητα. Αυτό απαιτεί τη σωστή και λεπτομερή διατύπωσή του, εστίαση σε πραγματικά υπάρχοντα προβλήματα, και τον καθορισμό του πλαισίου του. Κατά τη διαδικασία αυτή, πρέπει να ορίζονται σαφώς οι στόχοι της ανάπτυξης και χρήσης του ΓΣΠ, καθώς και προκαταρκτικές ενέργειες για την εισαγωγή δεδομένων στον υπολογιστή.

3.2.2 Διαδικασία μετατροπής των στοιχείων σε πληροφορία

Η διαδικασία αυτή αποτελεί τον πυρήνα κάθε Γεωγραφικού Συστήματος Πληροφοριών (ΓΣΠ) και περιλαμβάνει τέσσερα βασικά στάδια:

1. Στάδιο Εισόδου

Σε αυτό το στάδιο πραγματοποιείται η κωδικοποίηση και αποθήκευση όλων των απαραίτητων δεδομένων, χωρικών και μη, στον υπολογιστή. Από την οπτική των Γεωγραφικών Συστημάτων Πληροφοριών (GIS), πρόκειται για τη διαδικασία συλλογής και καταγραφής των στοιχείων που απαιτούνται για κάθε εφαρμογή. Τα δεδομένα αυτά μπορεί να προέρχονται είτε από πρωτογενείς πηγές, όπως η θεωρητική μελέτη και η έρευνα πεδίου, είτε από την επεξεργασία ήδη υπάρχοντων δεδομένων, όπως η ψηφιοποίηση χαρτών, αλλά και από οργανωμένες βάσεις δεδομένων.

2. Σύστημα Διαχείρισης

Στο στάδιο αυτό, τα συλλεχθέντα χωρικά δεδομένα οργανώνονται και μετασχηματίζονται σε δομημένες πληροφοριακές βάσεις, προκειμένου να προετοιμαστούν για περαιτέρω ανάλυση. Η διαχείριση στα GIS περιλαμβάνει τη διάρθρωση και τη συστηματική οργάνωση των πληροφοριών που αφορούν τις γεωγραφικές οντότητες, όπως η θέση, η τοπολογία και τα χαρακτηριστικά τους. Η διαδικασία αυτή συνδέεται άμεσα με την έννοια του Συστήματος Διαχείρισης Δεδομένων και θεωρείται το θεμέλιο πάνω στο οποίο στηρίζεται η λειτουργία ενός GIS.

3. Στάδιο Ανάλυσης

Σε αυτό το στάδιο εφαρμόζονται κατάλληλες μεθόδους ανάλυσης επί των χωρικών δεδομένων. Οι τεχνικές που χρησιμοποιούνται δεν είναι σταθερές, καθώς εξαρτώνται τόσο από τον στόχο της ανάλυσης όσο και από τα χαρακτηριστικά των δεδομένων και το διαθέσιμο τεχνολογικό υπόβαθρο. Στο πλαίσιο της χωρικής ανάλυσης, τα ΓΣΠ απαντούν σε πέντε βασικούς τύπους ερωτημάτων:

- "Τι βρίσκεται πού;", όπου προσδιορίζεται τι υπάρχει σε μια συγκεκριμένη θέση βάσει συντεταγμένων ή άλλων γεωγραφικών στοιχείων.
- Πού βρίσκεται κάτι;", με σκοπό τον εντοπισμό χαρακτηριστικών σε συγκεκριμένες θέσεις.
- "Τι άλλαξε;", μέσω της σύγκρισης στοιχείων σε διαφορετικές χρονικές στιγμές για την αναγνώριση μεταβολών.
- "Ποια χωρικά πρότυπα υπάρχουν;", όπου αναζητούνται χωρικές σχέσεις και κανόνες που διέπουν φαινόμενα.
- "Τι θα συμβεί αν;", που απαιτούν συνδυασμό γεωγραφικής πληροφορίας και μοντέλων πρόβλεψης ή προσομοίωσης για την ανάλυση πολύπλοκων φαινομένων.

4. Στάδιο Εξόδου

Σε αυτό το στάδιο, τα αποτελέσματα της ανάλυσης παρουσιάζονται σε μορφές όπως πίνακες, γραφήματα και κυρίως χάρτες. Οι βασικές μορφές απεικόνισης περιλαμβάνουν: (α) πίνακες και μαθηματικά μοντέλα, (β) γραφήματα όπως ιστογράμματα και πολύγωνα συχνότητας και (γ) χαρτογραφικές αποδόσεις. Τα αποτελέσματα μπορούν να προβληθούν στην οθόνη του υπολογιστή, να αποθηκευτούν, να εκτυπωθούν ή να διαμοιραστούν διαδικτυακά, ενώ χρησιμοποιούνται επίσης ως βάση για την ανάπτυξη εφαρμογών. Ο χάρτης αποτελεί το βασικό μέσο αποτύπωσης της γεωγραφικής πληροφορίας, προσφέροντας εποπτεία και κατανόηση των χωρικών δεδομένων με υψηλή ακρίβεια και σαφήνεια. Ιδιαίτερη σημασία έχουν οι θεματικοί χάρτες, οι οποίοι διευκολύνουν την οπτική κατανόηση φυσικών ή ανθρωπογενών φαινομένων, καθώς οι άνθρωποι αντιλαμβάνονται ευκολότερα πληροφορίες που αποδίδονται οπτικά παρά μέσω αριθμητικών δεδομένων (Κουτσόπουλος Κ., 2002).

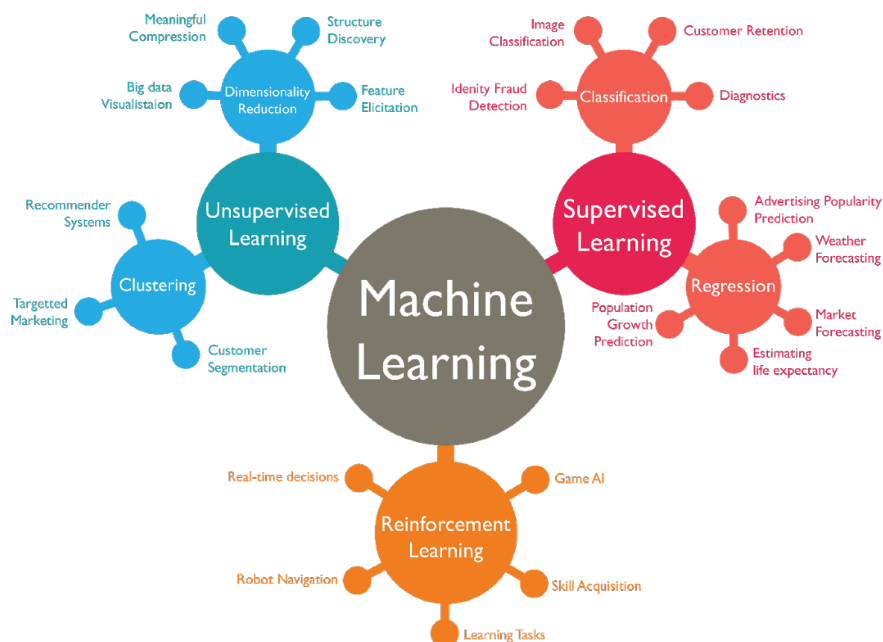
3.2.3 Διαδικασία εξαγωγής συμπερασμάτων

Στο στάδιο της εξαγωγής συμπερασμάτων του ΓΣΠ, το βασικό ζητούμενο είναι η απόδοση της πληροφορίας που προκύπτει από τη διαδικασία ανάλυσης, με σκοπό την επίλυση του προβλήματος. Είναι απαραίτητο να αντικατοπτρίζεται τόσο η επίτευξη του στόχου της χρήσης του ΓΣΠ, όσο και η ολοκληρωμένη χωρική προσέγγιση και οι εναλλακτικές απόψεις για την αντιμετώπιση του προβλήματος. Οι προτεινόμενες λύσεις παρουσιάζονται με στόχο να απαντηθούν ερωτήματα όπως "Τί είναι; Τί πρέπει; Τί είναι εφικτό;". Τα συμπεράσματα πρέπει να τεκμηριώνονται με κατάλληλα στοιχεία, ιδίως όταν βασίζονται σε γνώμη ή κρίση του ερευνητή. Επίσης, είναι απαραίτητο να διευκρινίζονται οι λόγοι για τους οποίους καταλήγουμε στη συγκεκριμένη κρίση, να παρέχεται δηλαδή τεκμηρίωση τόσο για τη διαδικασία όσο και για το αποτέλεσμα. Στο πλαίσιο της ολοκληρωμένης χωρικής προσέγγισης, η

ανάλυση μέσω Γεωγραφικών Συστημάτων Πληροφοριών (ΓΣΠ) αποτελεί θεμελιώδες στάδιο για την εξαγωγή συμπερασμάτων, τα οποία δύνανται είτε να υποστηρίξουν καθαρά ερευνητικούς στόχους είτε να συμβάλουν στη λήψη αποφάσεων για χωρικές παρεμβάσεις. Ειδικότερα, τα ΓΣΠ δεν περιορίζονται αποκλειστικά στην υποστήριξη επεμβάσεων στον χώρο, αλλά χρησιμοποιούνται ευρέως για τη μελέτη, την κατανόηση και τη μοντελοποίηση γεωγραφικών φαινομένων (όπως η ανάλυση της σχέσης μεταξύ βροχοπτώσεων και βλάστησης). Σε περιπτώσεις που απαιτείται σχεδιασμός και υλοποίηση επεμβάσεων, τα ΓΣΠ διαδραματίζουν κρίσιμο ρόλο τόσο στη συλλογή και επεξεργασία του αναγκαίου χωρικού υλικού όσο και στην υποστήριξη της επιλογής της βέλτιστης μεθοδολογίας παρέμβασης, διασφαλίζοντας τη συνεκτικότητα και τη λειτουργικότητα του χώρου.

4 Μηχανική Μάθηση

Η μηχανική μάθηση (Machine Learning - ML) είναι ένας τομέας της τεχνητής νοημοσύνης που επικεντρώνεται στη δημιουργία συστημάτων ικανών να μαθαίνουν και να βελτιώνονται μέσω της εμπειρίας. Αυτά τα συστήματα χρησιμοποιούν προηγούμενα δεδομένα για να εξελίσσονται και να αποδίδουν καλύτερα σε συγκεκριμένες διαδικασίες, χωρίς να χρειάζεται περαιτέρω προγραμματισμός από τον άνθρωπο. Η βελτίωσή τους επιτυγχάνεται αυτόματα, καθώς αντλούν γνώση από τα δεδομένα και τις εμπειρίες που αποκτούν κατά τη λειτουργία τους.



Εικόνα 3: Είδη και Εφαρμογές της Μηχανικής Μάθησης¹

4.1 Είδη Μηχανικής Μάθησης

Η Μηχανική Μάθηση χωρίζεται σε τρεις κατηγορίες ανάλογα με τον τύπο ανάδρασης. Οι κατηγορίες αυτές είναι η επιβλεπόμενη μάθηση (supervised machine learning), η μη επιβλεπόμενη μάθηση (unsupervised machine learning) και η ενισχυτική μάθηση (reinforcement learning).

4.1.1 Επιβλεπόμενη Μηχανική Μάθηση

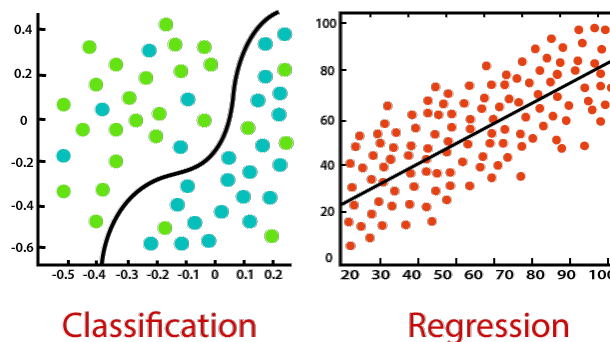
Η επιβλεπόμενη μηχανική μάθηση είναι μια από τις πιο δημοφιλείς και αποτελεσματικές προσεγγίσεις στη μηχανική μάθηση. Χρησιμοποιείται για την πρόβλεψη ενός συγκεκριμένου αποτελέσματος, βασιζόμενη σε ένα προκαθορισμένο σύνολο δεδομένων εκπαίδευσης. Ονομάζεται "επιβλεπόμενη" στα δεδομένα εκπαίδευσης που παρέχουμε στον αλγόριθμο έχουν αντιστοιχηθεί «ετικέτες» (labels) οι οποίες αντανακλούν τις κατηγορίες στις οποίες εντάσσονται τα δεδομένα τιμές κάποιων χαρακτηριστικών τους. Μετά την εκπαίδευση, δημιουργείται ένα μοντέλο που μπορεί να εφαρμοστεί σε νέα δεδομένα για να κάνει προβλέψεις. Υπάρχουν δύο βασικά είδη προβλημάτων στα οποία εφαρμόζεται αυτή η μέθοδος:

¹ Πηγή: <https://7wdata.be/visualization/types-of-machine-learning-algorithms-2/>

- Κατηγοριοποίηση, όπου το ζητούμενο είναι η ταξινόμηση των δεδομένων σε κατηγορίες (π.χ., "ναι" ή "όχι", "γάτα" ή "σκύλος").
- Παλινδρόμηση, όπου το ζητούμενο είναι η πρόβλεψη συνεχών τιμών (π.χ., η τιμή ενός σπιτιού ή η θερμοκρασία).

Στην παρούσα εργασία, η ανάλυση εστιάζει σε δεδομένα από τηλεπισκοπικές απεικονίσεις, όπου τα εικονοστοιχεία (pixels) αντιπροσωπεύουν τιμές φασματικών καναλιών. Τα χαρακτηριστικά των εικονοστοιχείων χρησιμοποιούνται για την κατηγοριοποίηση ή την παλινδρόμηση με βάση φασματικές υπογραφές αναφοράς.

Στις επόμενες παραγράφους αναλύονται ορισμένα από τα πιο σημαντικά χρησιμοποιούμενα είδη ταξινόμησης και παλινδρόμησης.



Εικόνα 4: Κατηγοριοποίηση - Παλινδρόμηση²

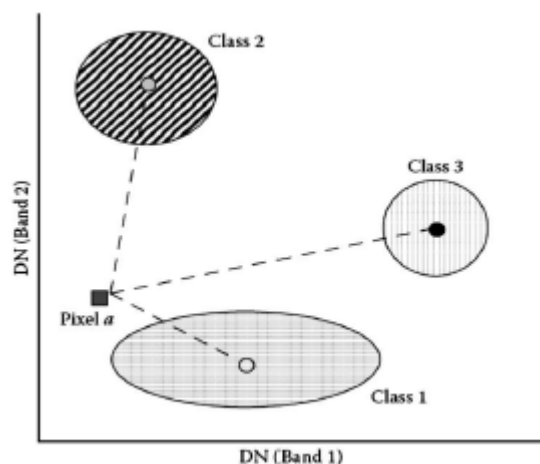
Η κατηγοριοποίηση (classification) αποτελεί κεντρική διαδικασία της επιβλεπόμενης μηχανικής μάθησης, όπου ο στόχος είναι η αντιστοίχιση παρατηρήσεων σε προκαθορισμένες κατηγορίες (Géron, A, 2022). Σύγχρονες τεχνικές περιλαμβάνουν αλγορίθμους όπως τα δέντρα αποφάσεων, τα τυχαία δάση (Random Forests) και τις μηχανές διανυσμάτων υποστήριξης (SVMs), που μπορούν να χειριστούν τόσο γραμμικά όσο και μη γραμμικά διαχωρίσιμα δεδομένα (Raschka & Mirjalili, 2022). Στον τομέα της τηλεπισκόπησης, η κατηγοριοποίηση χρησιμοποιείται ευρέως για την ανάλυση φασματικών χαρακτηριστικών και την ταξινόμηση εικονοστοιχείων σε τύπους κάλυψης γης. Η ακρίβεια των μοντέλων αξιολογείται με δείκτες όπως η συνολική ακρίβεια και οι πίνακες σύγχυσης (confusion matrices).

4.1.1.1 Κατηγοριοποίηση Μέγιστης Πιθανοφάνειας (Maximum Likelihood)

Στην Κατηγοριοποίηση Μέγιστης Πιθανοφάνειας υπολογίζεται η διασπορά και η συνδιασπορά της κάθε κατηγορίας όταν γίνεται κατηγοριοποίηση ενός αγνώστου εικονοστοιχείου. Η υπόθεση που γίνεται είναι ότι σε κάθε κατηγορία, η κατανομή του νέφους των σημείων (διανύσματα τιμών των εικονοστοιχείων) που απαρτίζουν την κατηγορία είναι κανονική (Gaussian, normally distributed). Κάτω από αυτή την υπόθεση η κατανομή της φασματικής απόκρισης της τάξης, δηλαδή η κατανομή των τιμών ανακλαστικότητας ή φασματικών χαρακτηριστικών των εικονοστοιχείων που ανήκουν σε μία δεδομένη κατηγορία, όπως αυτή παρατηρείται σε πολλαπλά φασματικά κανάλια, μπορεί να περιγράψει πλήρως

² Πηγή: <https://www.tpointtech.com/regression-vs-classification-in-machine-learning>

από το μέσο διάνυσμα και τον πίνακα συνδιασποράς. Με δεδομένες τις παραπάνω παραμέτρους μπορεί να υπολογιστεί η στατιστική πιθανότητα να ανήκει ένα εικονοστοιχείο σε μία κατηγορία.

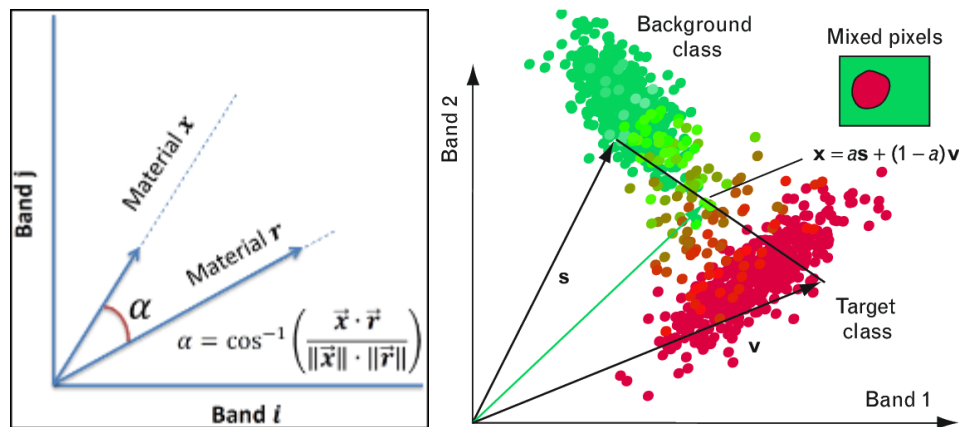


Εικόνα 5: Ταξινόμηση Μέγιστης Πιθανοφάνειας³

4.1.1.2 Κατηγοριοποίηση Φασματικής Γωνίας (Spectral Angle Mapping)

Η κατηγοριοποίηση φασματικής γωνίας βασίζεται στη φυσική του φάσματος και χρησιμοποιεί τη Ν-διάστατη γωνία για να εντάξει κάποια εικονοστοιχεία σε μία φασματική υπογραφή αναφοράς. Ο αλγόριθμος προσδιορίζει την ομοιότητα μεταξύ δύο φασματικών υπογραφών υπολογίζοντας τη φασματική γωνία μεταξύ των δύο φασματικών υπογραφών, τις οποίες θεωρεί ως διανύσματα σε ένα χώρο με πλήθος διαστάσεων ίσο με το πλήθος καναλιών της εικόνας. Η τεχνική αυτή εφαρμόζεται συνήθως σε ατμοσφαιρικά διορθωμένες και βαθμονομημένες τηλεπισκοπικές απεικονίσεις, των οποίων οι ψηφιακές τιμές είναι τιμές ανακλαστικότητας και οι φασματικές υπογραφές αναφοράς μπορεί να προέρχονται είτε από φασματικές μετρήσεις πεδίου, είτε από φασματικές βιβλιοθήκες ή και από περιοχές εκπαίδευσης. Περιοχές εκπαίδευσης είναι επιλεγμένες περιοχές της τηλεπισκοπικής εικόνας όπου η φυσική ταυτότητα κατανομής είναι γνωστή. Χρησιμοποιούνται για εκπαίδευση αλγορίθμων (Richards & Jia, 2006). Οι φασματικές βιβλιοθήκες (Spectral Libraries) είναι συλλογές από φασματικές υπογραφές (spectral signatures) διαφορετικών υλικών (π.χ. εδάφη, βλάστηση, νερά, ορυκτά), οι οποίες έχουν καταγραφεί είτε με χρήση φασματογράφων πεδίου είτε από δορυφορικές/εναέριες λήψεις. (Clark 2023) Αυτές οι υπογραφές χρησιμοποιούνται ως σημεία αναφοράς για την κατηγοριοποίηση ή ταυτοποίηση άγνωστων εικονοστοιχείων σε τηλεπισκοπικές εικόνες. Η μέθοδος αυτή συγκρίνει τις φασματικές γωνίες μεταξύ των φασματικών υπογραφών αναφοράς και το φασματικό διάνυσμα κάθε εικονοστοιχείου σε ένα χώρο Ν διαστάσεων (όπου Ν ο αριθμός των καναλιών). Μικρές γωνίες συνεπάγονται μεγάλη ομοιότητα ενώ μεγάλες γωνίες συνεπάγονται μικρή ομοιότητα.

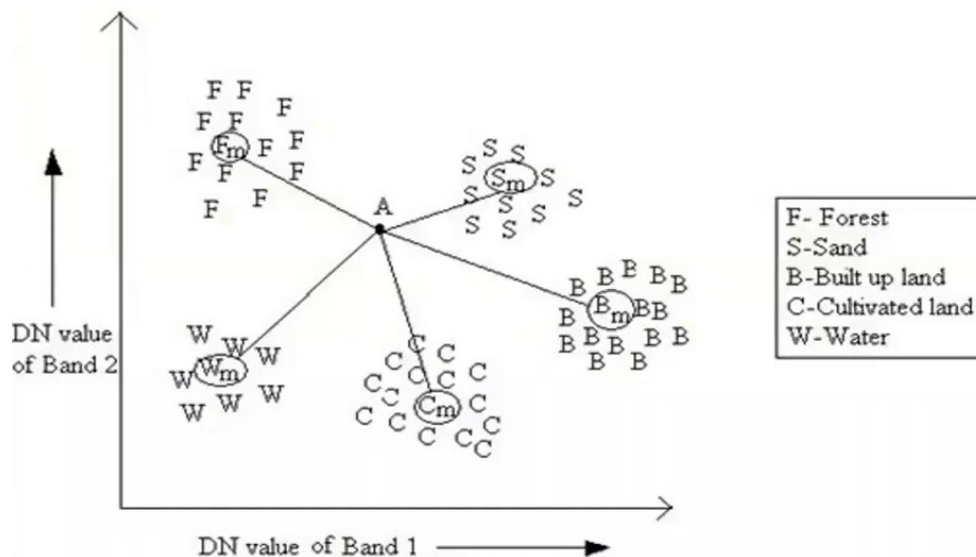
³ Πηγή: Gao J. (2009). Digital Analysis of Remote Sensing Imagery, McGraw-Hill Companies, Inc.



Εικόνα 6: Ταξινόμηση Φασματικής Γωνίας⁴

4.1.1.3 Κατηγοροποίηση Ελάχιστης Απόστασης (Minimum Distance)

Είναι μία από τις απλούστερες μεθοδολογίες επιβλεπόμενης κατηγοριοποίησης. Σε αυτήν υπολογίζονται οι μέσες τιμές για κάθε φασματικό κανάλι των περιοχών εκπαίδευσης. Αυτές οι τιμές αντιπροσωπεύουν το διάνυσμα των μέσων τιμών των περιοχών εκπαίδευσης (καθώς και των αντίστοιχων κατηγοριών). Κάθε εικονοστοιχείο εντάσσεται σε μία κατηγορία με βάση την απόσταση του από το διάνυσμα μέσων τιμών της κατηγορίας, δηλαδή εντάσσεται στην κατηγορία στην οποία η απόσταση της αντίστοιχης περιοχής εκπαίδευσης από την προβολή του εικονοστοιχείου στο φασματικό χώρο προτύπων (που ορίζεται από τα φασματικά κανάλια) είναι ελάχιστη.



Εικόνα 7: Ταξινόμηση Ελάχιστης Απόστασης⁵

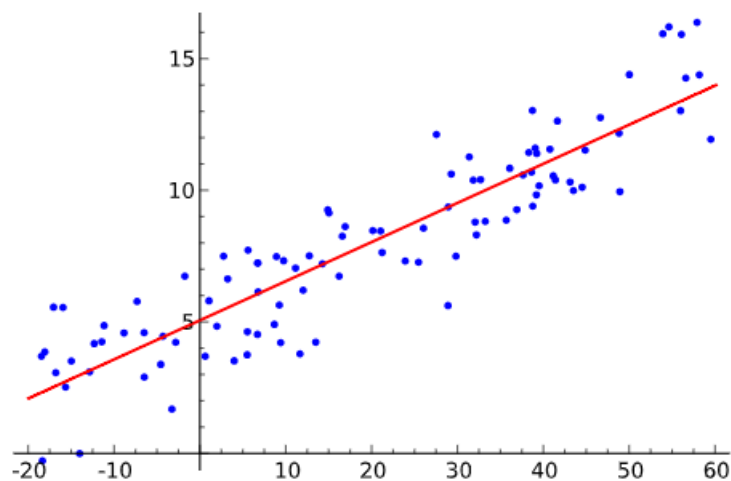
⁴ Πηγή: Manolakis, D., Marden, D., & Shaw, G. A. (2003). Hyperspectral Image Processing for Automatic Target Detection Applications. LINCOLN LABORATORY JOURNAL, 14(1).

⁵ Πηγή: <https://www.youtube.com/watch?v=49OHn4WcX1Y>

Η παλινδρόμηση αποτελεί βασικό κλάδο της επιβλεπόμενης μηχανικής μάθησης και χρησιμοποιείται για την πρόβλεψη συνεχών ποσοτικών μεταβλητών (Géron, 2022). Σε αντίθεση με την κατηγοριοποίηση, όπου προβλέπονται διακριτές ετικέτες, η παλινδρόμηση έχει ως στόχο την εκτίμηση μίας τιμής που κυμαίνεται σε συνεχή αριθμητική κλίμακα. Οι εφαρμογές της περιλαμβάνουν, μεταξύ άλλων, την πρόβλεψη τιμών ακινήτων, θερμοκρασιών, συγκεντρώσεων ρύπων ή γεωφυσικών μετρήσεων, και έχουν ιδιαίτερη σημασία και σε περιβάλλοντα ανάλυσης τηλεπισκοπικών δεδομένων, όπου η εκτίμηση συνεχών παραμέτρων από εικόνες αποτελεί κεντρική ανάγκη.

4.1.1.4 Γραμμική Παλινδρόμηση (Linear Regression)

Η γραμμική παλινδρόμηση (Linear Regression) είναι από τις παλαιότερες και πιο θεμελιώδεις τεχνικές πρόβλεψης (Bishop, 2006). Το βασικό της μοντέλο στηρίζεται στην υπόθεση ότι υπάρχει γραμμική σχέση μεταξύ των εισόδων και της εξόδου. Ο αλγόριθμος αναζητά την ευθεία που ελαχιστοποιεί το άθροισμα των τετραγώνων των διαφορών μεταξύ των προβλεπόμενων και των πραγματικών τιμών (μέθοδος ελαχίστων τετραγώνων). Παρά την απλότητά της, η γραμμική παλινδρόμηση συχνά λειτουργεί εξαιρετικά καλά σε προβλήματα όπου οι σχέσεις είναι πράγματι γραμμικές ή προσεγγίζονται επαρκώς ως τέτοιες. Στη σύγχρονη πράξη χρησιμοποιούνται και επεκτάσεις όπως η Ρυθμισμένη Γραμμική Παλινδρόμηση (Regularized Linear Regression, π.χ. Ridge ή Lasso) για την αποφυγή υπερπροσαρμογής (overfitting) (Géron, A, 2022).



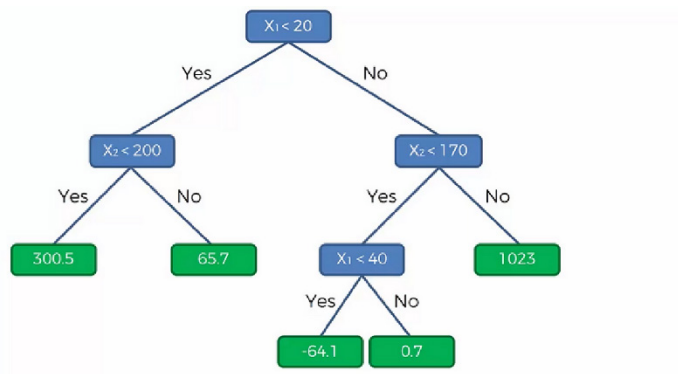
Εικόνα 8: Γραμμική Παλινδρόμηση⁶

4.1.1.5 Δέντρα Αποφάσεων για Παλινδρόμηση (Decision Trees Regression)

Τα δέντρα αποφάσεων (Decision Trees) χρησιμοποιούνται εκτενώς για παλινδρόμηση όταν η σχέση μεταξύ μεταβλητών είναι μη γραμμική ή περιπλοκή (Goodfellow et al., 2016). Ένα δέντρο αποφάσεων δημιουργεί διαδοχικές διαιρέσεις του χώρου εισόδων βασισμένες σε χαρακτηριστικά τιμών, επιδιώκοντας σε κάθε κόμβο να μειώσει το σφάλμα παλινδρόμησης (π.χ. μέσο τετραγωνικό σφάλμα - MSE). Το μοντέλο είναι εύκολα ερμηνεύσιμο και μπορεί να μοντελοποιήσει πολύπλοκες, μη γραμμικές

⁶ Géron, A. – Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (3η έκδοση, 2022)

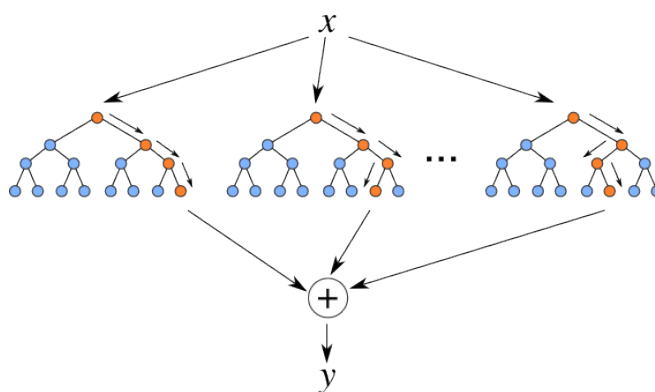
σχέσεις, αν και συχνά κινδυνεύει από υπερπροσαρμογή, ειδικά αν δεν περιοριστεί το βάθος του δέντρου (Géron, 2022).



Εικόνα 9: Δέντρα Αποφάσεων για Παλινδρόμηση⁷

4.1.1.6 Παλινδρόμηση με Random Forests (Random Forest Regression)

Το Random Forest Regression αποτελεί βελτίωση της μεθόδου των δέντρων αποφάσεων, συνδυάζοντας πολλά ανεξάρτητα δέντρα σε ένα ensemble (Breiman, 2001). Κάθε δέντρο εκπαιδεύεται σε διαφορετικό δείγμα των δεδομένων με τη μέθοδο bootstrap και η τελική πρόβλεψη είναι ο μέσος όρος των προβλέψεων όλων των δέντρων. Αυτή η τεχνική μειώνει σημαντικά τον κίνδυνο υπερπροσαρμογής που χαρακτηρίζει τα μεμονωμένα δέντρα και προσφέρει υψηλή ακρίβεια σε ένα ευρύ φάσμα προβλημάτων παλινδρόμησης. Επιπλέον, το Random Forest παρέχει τη δυνατότητα εκτίμησης της σχετικής σημασίας κάθε χαρακτηριστικού στα δεδομένα (Géron, 2022).



Εικόνα 10: Random Forest⁸

4.1.2 Μη επιβλεπόμενη Μηχανική Μάθηση

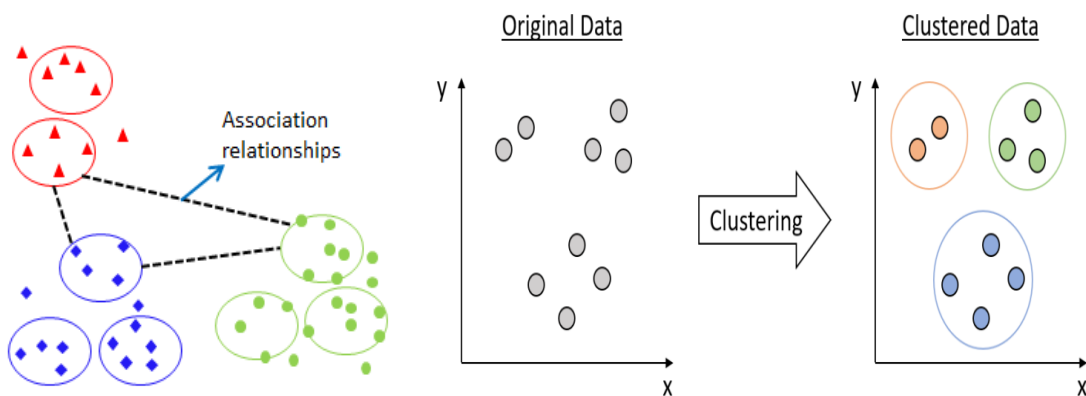
Η μη επιβλεπόμενη μάθηση είναι ένας άλλος τύπος μηχανικής μάθησης. Σε αντίθεση με την επιβλεπόμενη μάθηση, όπου καθορίζουμε τις πιθανές τιμές που μπορεί να έχει η κατηγορία-στόχος, εδώ δίνονται μόνο τα δεδομένα εισόδου χωρίς να έχουν αντιστοιχηθεί σε αυτά ετικέτες (labels). Ουσιαστικά, ο αλγόριθμος παίρνει αυτά τα δεδομένα και προσπαθεί να βγάλει νόημα από μόνος του, χωρίς να του

⁷ Πηγή: <https://maniksonituts.medium.com/what-is-decision-tree-regression-dcd0ea40a323>

⁸ Πηγή: <https://medium.com/@vk.viswa/surface-level-understanding-of-random-forest-regression-3bac37e06f3e>

δώσουμε εμείς κάποια καθοδήγηση. Παρόλο που υπάρχουν πολλές επιτυχημένες εφαρμογές αυτής της προσέγγισης, είναι συνήθως πιο δύσκολο να καταλάβουμε, να ερμηνεύσουμε, και να αξιολογήσουμε τα αποτελέσματα που προκύπτουν. Η μη επιβλεπόμενη μάθηση χωρίζεται σε δύο βασικές κατηγορίες:

- Μετασχηματισμός Δεδομένων: Περιλαμβάνει τεχνικές όπως η μείωση διαστάσεων (π.χ., PCA) για την απλοποίηση και καλύτερη οργάνωση των δεδομένων.
- Συσταδοποίηση (Clustering): Ομαδοποιεί τα δεδομένα με βάση τις ομοιότητες τους, ώστε να εντοπιστούν πρότυπα ή δομές.

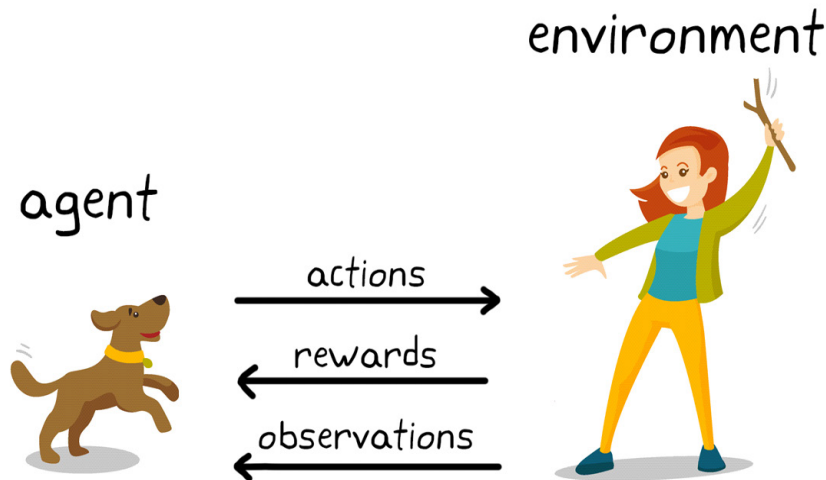


Εικόνα 11. Μη επιβλεπόμενη μάθηση⁹

4.1.3 Ενισχυτική Μάθηση

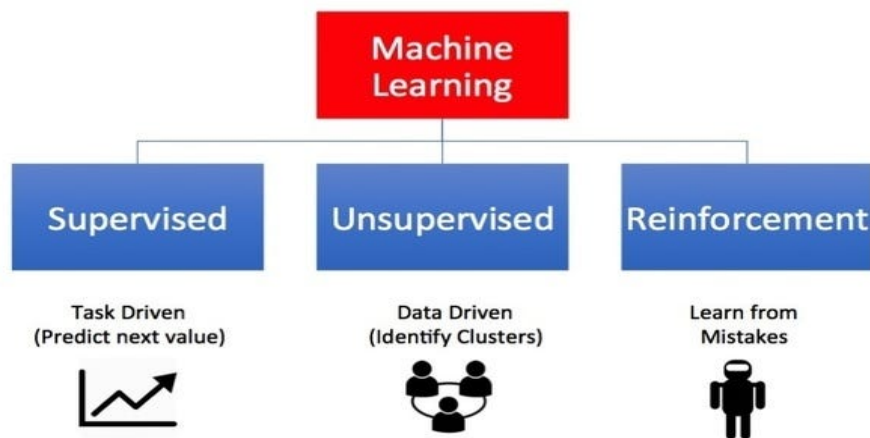
Η ενισχυτική μάθηση αποτελεί το τρίτο είδος μηχανικής μάθησης, λιγότερο διαδεδομένο σε σχέση με την επιβλεπόμενη και τη μη επιβλεπόμενη μάθηση. Σε αυτό το είδος, το σύστημα εκπαιδεύεται μέσω αλληλεπίδρασης με ένα αβέβαιο περιβάλλον. Χρησιμοποιώντας τη μέθοδο της δοκιμής και σφάλματος, προσπαθεί να ανακαλύψει την καλύτερη δυνατή στρατηγική για την επίτευξη του επιθυμητού αποτελέσματος. Στην ουσία, διερευνάται το πώς ένας αυτόνομος πράκτορας, ο οποίος μπορεί να αισθάνεται και να ενεργεί στο περιβάλλον του, μαθαίνει να επιλέγει τις κατάλληλες ενέργειες που θα τον οδηγήσουν στην επίτευξη των στόχων του.

⁹ Πηγή: K. Luong and R. Nayak, "Learning Association Relationship and Accurate Geometric Structures for Multi-Type Relational Data," *2018 IEEE 34th International Conference on Data Engineering (ICDE)*, Paris, France, 2018, pp. 509-520, doi: 10.1109/ICDE.2018.00053.



Εικόνα 12: Ενισχυτική Μάθηση¹⁰

Types of Machine Learning



Εικόνα 13: Είδη Μηχανικής Μάθησης¹¹

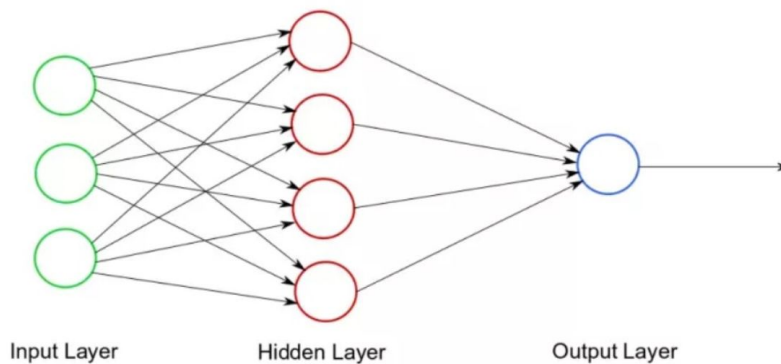
4.2 Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα αποτελούν ένα θεμελιώδες εργαλείο (Géron, A. (2022) στη μηχανική μάθηση και ιδιαίτερα στη βαθιά μάθηση (deep learning). Εμπνευσμένα από τη δομή και τη λειτουργία του ανθρώπινου εγκεφάλου, αποτελούνται από τεχνητούς «νευρώνες», οι οποίοι είναι διασυνδεδεμένοι μεταξύ τους σχηματίζοντας ένα δίκτυο. Κάθε νευρώνας λαμβάνει δεδομένα ως είσοδο, εκτελεί έναν υπολογισμό μέσω μιας συνάρτησης ενεργοποίησης, και στη συνέχεια προωθεί το αποτέλεσμα στους επόμενους νευρώνες. Το δίκτυο περιλαμβάνει τρία βασικά είδη στρωμάτων: το εισόδου (input layer), που δέχεται τα δεδομένα, το κρυφό (hidden layer), όπου γίνεται η επεξεργασία μέσω πολλαπλών

¹⁰ Πηγή: <https://onlineacademi.wordpress.com/2019/12/16/what-are-machine-learning-algorithms-and-its-different-types/>

¹¹ Πηγή: AnalyticSteps.com

επιπέδων, και το εξόδου (output layer), που παράγει το τελικό αποτέλεσμα. Η εκπαίδευση ενός νευρωνικού δικτύου γίνεται μέσω ενός αλγορίθμου βελτιστοποίησης, που προσαρμόζει τα βάρη των συνδέσεων μεταξύ των νευρώνων ώστε να ελαχιστοποιηθεί το σφάλμα στις προβλέψεις. Τα νευρωνικά δίκτυα έχουν τη δυνατότητα να μαθαίνουν από πολύπλοκα, μη γραμμικά δεδομένα και να αναγνωρίζουν πρότυπα που δύσκολα ανιχνεύονται με παραδοσιακές μεθόδους, γεγονός που τα καθιστά ιδανικά για προβλήματα υψηλής πολυπλοκότητας και μεγάλης κλίμακας δεδομένων. Ωστόσο, η αποτελεσματική χρήση τους απαιτεί μεγάλη υπολογιστική ισχύ, επαρκή δεδομένα για εκπαίδευση, και προσεκτικό σχεδιασμό της αρχιτεκτονικής τους, καθώς ζητήματα όπως η υπερπροσαρμογή και η επιλογή υπερπαραμέτρων μπορούν να επηρεάσουν σημαντικά την απόδοσή τους. Εν τέλει, τα νευρωνικά δίκτυα δεν είναι απλά ένας αλγόριθμος αλλά ένα εργαλείο που αναδεικνύει τη δυναμική της τεχνητής νοημοσύνης για την επίλυση πολύπλοκων προβλημάτων σε ποικίλα πεδία εφαρμογών.



Εικόνα 14: Νευρωνικό Δίκτυο

5 Τηλεπισκόπηση

Η δορυφορική τηλεπισκόπηση είναι η επιστήμη και η τεχνολογία που επιτρέπει τη συλλογή δεδομένων για έναν χώρο ή ένα φαινόμενο χωρίς την ανάγκη φυσικής επαφής. Ο όρος προέρχεται από τις λέξεις «τήλε» (από απόσταση) και «επισκοπώ» (παρατηρώ οπτικά). Η μέθοδος αυτή εφαρμόζεται εκτενώς στην παρατήρηση της επιφάνειας της Γης. Μέσα από την αλληλεπίδραση των υλικών της Γης με την ηλεκτρομαγνητική ακτινοβολία, συλλέγονται και αναλύονται πληροφορίες από διαφορετικά μέρη του ηλεκτρομαγνητικού φάσματος. Η τηλεπισκόπηση βρίσκει εφαρμογές σε τομείς όπως η μετεωρολογία, η μελέτη του περιβάλλοντος, η χαρτογραφία, η πρόβλεψη και η διαχείριση φυσικών καταστροφών, η γεωργία, και η παρακολούθηση του κλίματος και των ανθρωπογενών επιπτώσεων στο περιβάλλον.

Στη συνέχεια περιγράφονται δύο προγράμματα τηλεπισκόπησης μέσω δορυφόρων, τα οποία συλλέγουν και διαθέτουν δεδομένα που θα χρησιμοποιηθούν στην παρούσα εργασία.

5.1 Πρόγραμμα Copernicus

Το πρόγραμμα Copernicus (<https://www.copernicus.eu/>) έχει σχεδιαστεί για τη συγκέντρωση και ανάλυση μεγάλου όγκου γεωπαρατηρησιακών δεδομένων, με εφαρμογές που καλύπτουν την κλιματική αλλαγή, τη βιώσιμη ανάπτυξη, τη διαχείριση φυσικών πόρων, καθώς και τη μελέτη της ατμόσφαιρας, της επιφάνειας της Γης, του θαλάσσιου περιβάλλοντος, αλλά και την αντιμετώπιση εκτάκτων καταστάσεων. Στο πλαίσιο του προγράμματος, δημιουργήθηκε η οικογένεια δορυφόρων Sentinel, με εξειδικευμένες αποστολές:

- Sentinel-1: Παρατήρηση χερσαίων και θαλάσσιων περιοχών με ραντάρ ανεξάρτητο από καιρικές συνθήκες.
- Sentinel-2: Πολυφασματικές εικόνες υψηλής ανάλυσης για μελέτες βλάστησης, υδάτινων σωμάτων και κάλυψης εδάφους.
- Sentinel-3: Δεδομένα τοπογραφίας, θερμοκρασίας και χρωματισμού ωκεανών και ξηράς.
- Sentinel-4 και Sentinel-5/5P: Μελέτη ατμοσφαιρικής σύνθεσης, παρακολούθηση ποιότητας αέρα και ηλιακής ακτινοβολίας.
- Sentinel-6: Υψομετρικά δεδομένα για την παρακολούθηση της στάθμης της θάλασσας και των κλιματικών αλλαγών.

Επιπλέον, η τηλεπισκόπηση επεκτείνεται με την ανάπτυξη μη επανδρωμένων εναέριων συστημάτων (UAVs/drones), τη χρήση τεχνητής νοημοσύνης και τη διασύνδεση δεδομένων σε πραγματικό χρόνο, αυξάνοντας τις δυνατότητες εφαρμογής και την ακρίβεια στην παρατήρηση και ανάλυση της Γης.

5.2 Sentinel 2

Η αποστολή Sentinel-2 περιλαμβάνει δύο δορυφόρους με πανομοιότυπα χαρακτηριστικά, οι οποίοι κινούνται στην ίδια ηλιοσύγχρονη τροχιά με διαφορά φάσης 180°. Ο Sentinel-2A εκτοξεύθηκε στις 23 Ιουνίου 2015 και ο Sentinel-2B στις 7 Μαρτίου 2017. Για να διασφαλιστεί η συνέχεια των παρατηρήσεων της Γης στο πλαίσιο του προγράμματος Copernicus, ο Sentinel-2C εκτοξεύθηκε στις 5 Σεπτεμβρίου 2024 με τον τελευταίο πύραυλο Vega, ενώ ο Sentinel-2D προγραμματίζεται να εκτοξευθεί το 2028. Κάθε

δορυφόρος έχει ονομαστική διάρκεια ζωής 7,25 χρόνια, ωστόσο με την επάρκεια καυσίμων και αναλωσίμων η λειτουργία του μπορεί να παραταθεί έως και 12 χρόνια.

Οι δορυφόροι Sentinel-2 κινούνται σε ηλιοσύγχρονη τροχιά σε υψόμετρο περίπου 786 χιλιομέτρων. Ο μέσος τοπικός ηλιακός χρόνος διέλευσης από τον καθοδικό κόμβο της τροχιάς είναι 10:30 π.μ., επιλογή που εξασφαλίζει καλό φωτισμό και μειώνει την πιθανότητα νεφοκάλυψης. Όταν λειτουργούν δύο δορυφόροι ταυτόχρονα, η αποστολή επιτυγχάνει χρονική διακριτική ικανότητα πέντε ημερών, προσφέροντας συχνές και επίκαιρες καταγραφές της ίδιας περιοχής της Γης. Ένας δορυφόρος ολοκληρώνει 143 τροχιές για την πλήρη κάλυψη της Γης, με κάθε τροχιά να διαρκεί περίπου 100,6 λεπτά.

Κάθε Sentinel-2 είναι εξοπλισμένος με τον πολυφασματικό αισθητήρα MSI (Multispectral Instrument), ο οποίος λειτουργεί με τη μέθοδο σάρωσης τύπου push-broom. Ο αισθητήρας διαθέτει πεδίο καταγραφής 290 χιλιομέτρων και χρόνο παρατήρησης που κυμαίνεται από 17 έως 32 λεπτά ανά πέρασμα. Η συλλογή της ακτινοβολίας πραγματοποιείται μέσω ενός τηλεσκοπίου τριών καθρεπτών, το οποίο διαχωρίζει την ακτινοβολία σε δύο εστιακά επίπεδα: το ένα για τις ζώνες του ορατού και του εγγύς υπέρυθρου (VNIR) και το άλλο για το υπέρυθρο βραχέος κύματος (SWIR), με χρήση διχρωικών διαχωριστών δέσμης. Ο περαιτέρω διαχωρισμός των φασματικών ζωνών επιτυγχάνεται μέσω φίλτρων ζωνών (stripe filters) τοποθετημένων πάνω στους ανιχνευτές.

Ο αισθητήρας MSI καταγράφει 13 φασματικές ζώνες που καλύπτουν περιοχές του ορατού, του εγγύς υπέρυθρου και του υπέρυθρου βραχέος κύματος. Παρέχει τρεις χωρικές αναλύσεις, 10m, 20m και 60m, ανάλογα με τη φασματική περιοχή, με τα δεδομένα 60m να χρησιμοποιούνται κυρίως για ατμοσφαιρικές διορθώσεις. Η φασματική διακριτική ικανότητα του αισθητήρα φτάνει τα 12 bit και το συνολικό βάρος του MSI είναι 290 κιλά. Τα δεδομένα που συλλέγονται από την αποστολή Sentinel-2 συμβάλλουν στην παρακολούθηση της γήινης επιφάνειας, στην εκτίμηση φυσικών καταστροφών, στη διαχείριση φυσικών πόρων, στην ασφάλεια και στην ανάλυση φαινομένων κλιματικής αλλαγής, ενισχύοντας τους στόχους του προγράμματος Copernicus. (ESA, 2024).

Ο Πίνακας 1 παρουσιάζει τα κανάλια του Sentinel 2 και τα χαρακτηριστικά τους.

Κανάλια Sentinel-2	Κεντρικό Μήκος Κύματος	Χωρική διακριτική ικανότητα (m)	Χρήση του καναλιού
1 Coastal aerosol	443	60	Παρακολούθηση παράκτιας ζώνης και ατμοσφαιρικών σωματιδίων (αερολύματα)
2 Blue	490	10	Παρακολούθηση υδάτινων σωμάτων, χαρτογράφηση φυτοπλαγκτού, διόρθωση ατμοσφαιρικών επιδράσεων
3 Green	560	10	Ανάλυση υγιούς βλάστησης, φυσικό χρωματικό φάσμα εικόνων
4 Red	665	10	Εντοπισμός βλάστησης, υγεία καλλιεργειών, χαρτογράφηση εδάφους
5 Vegetation Red Edge	705	20	Εκτίμηση υγείας και πυκνότητας βλάστησης
6 Vegetation Red Edge	740	20	Παρακολούθηση κατάστασης φυτικής κάλυψης και διαφοροποίηση ειδών βλάστησης
7 Vegetation Red Edge	783	20	Βελτιωμένη ανίχνευση βλάστησης και εφαρμογές γεωργίας ακριβείας
8 Εγγύς Υπέρυθρο NIR (ευρύ φάσμα)	842	10	Χαρτογράφηση βλάστησης, ανάλυση βιομάζας, υπολογισμός δεικτών βλάστησης όπως NDVI
8A Εγγύς Υπέρυθρο NIR (περιορισμένο φάσμα)	865	20	Βελτιωμένη εκτίμηση παραμέτρων βλάστησης και υγρασίας εδάφους
9 Water Vapour	945	60	Μέτρηση συγκέντρωσης υδρατμών στην ατμόσφαιρα
10 SWIR Ελικοειδής βλάστηση	1380	60	Ανίχνευση σύννεφων, υγρασίας και ατμοσφαιρικών χαρακτηριστικών
11 SWIR	1610	20	Χαρτογράφηση γεωλογικών χαρακτηριστικών, υγρασίας εδάφους και εκτίμηση πυρκαγιών
12 SWIR	2190	20	Εντοπισμός καμένων εκτάσεων, χαρακτηρισμός μεταλλευμάτων και εδαφικών τύπων

Πίνακας 1: Κανάλια Sentinel 2

6 Συνδυασμός GIS και Μηχανικής μάθησης

Οι σύγχρονες εξελίξεις στα Γεωγραφικά Πληροφοριακά Συστήματα (GIS), σε συνδυασμό με την ενσωμάτωση της γεωχωρικής τεχνητής νοημοσύνης (GeoAI), έχουν επιφέρει σημαντικές αλλαγές τόσο στην τεχνολογία όσο και στην κοινωνία. Η πρόοδος στον αυτοματισμό, σε συνδυασμό με τις καινοτομίες στον τομέα της τεχνητής νοημοσύνης, επιτρέπει την ανάλυση και εξαγωγή πολύτιμων πληροφοριών από μεγάλα και σύνθετα γεωγραφικά δεδομένα. Η αλληλεπίδραση του GeoAI με τη συνεχώς αυξανόμενη διαθεσιμότητα γεωγραφικών δεδομένων έχει ανοίξει νέους δρόμους σε πολλούς κλάδους. Από τη μελέτη των περιβαλλοντικών επιπτώσεων έως την ανάλυση τάσεων στις αγορές και τις κοινωνικές δυναμικές, το GeoAI συμβάλλει στην καλύτερη κατανόηση και διαχείριση κρίσιμων ζητημάτων. Παράλληλα, η αυτοματοποίηση μετασχηματίζει τον τρόπο λειτουργίας των επιχειρήσεων και της κοινωνίας γενικότερα, ενισχύοντας την αποδοτικότητα, την ακρίβεια και την ασφάλεια, ενώ περιορίζει την ανάγκη για ανθρώπινη παρέμβαση και τα πιθανά σφάλματα. Η επιρροή αυτών των τεχνολογικών εξελίξεων είναι εκτεταμένη, καλύπτοντας τομείς όπως η εκπαίδευση και η υγεία, δημιουργώντας νέες ευκαιρίες και βελτιώνοντας την καθημερινή ζωή. Στη συνέχεια, θα παρουσιαστούν ορισμένες από τις βασικές προκλήσεις που προκύπτουν από αυτές τις εξελίξεις (Shi, 2020).

Η ενοποίηση της μηχανικής μάθησης, της αναγνώρισης προτύπων και της αυτοματοποιημένης επεξεργασίας δεδομένων με τα ΓΠΣ διευκολύνει την ανάλυση γεωγραφικών περιοχών. Αυτή η συνδυαστική προσέγγιση επιτρέπει τον εντοπισμό μοτίβων στα δεδομένα, την πρόβλεψη μελλοντικών τάσεων και την αυτοματοποιημένη επεξεργασία εικόνων ή γεωχωρικών πληροφοριών. Η αξιοποίηση αυτών των τεχνικών συμβάλλει στην καλύτερη κατανόηση των γεωγραφικών συνθηκών, στην πρόβλεψη μεταβολών και στη λήψη τεκμηριωμένων αποφάσεων σε διάφορους τομείς, όπως ο πολεοδομικός σχεδιασμός, η προστασία του περιβάλλοντος και η διαχείριση φυσικών καταστροφών (Mukhamediev, Ravil I. 2022). Ο συνδυασμός της τεχνητής νοημοσύνης με τα ΓΠΣ ενισχύει και βελτιώνει την ανάλυση χωρικών δεδομένων. Το ΓΠΣ μπορεί να αξιοποιήσει τα αποτελέσματα που προκύπτουν από την αναγνώριση προτύπων μέσω AI και να τα επεξεργαστεί περαιτέρω, επιτρέποντας μια πιο λεπτομερή ανάλυση. Η ενσωμάτωση χωρικής οπτικοποίησης και ανάλυσης με τα ΓΠΣ παρέχει τη δυνατότητα παρουσίασης δεδομένων με τρόπο πιο σαφή και εύκολα κατανοητό, διευκολύνοντας τους χρήστες, όπως οι υπεύθυνοι λήψης αποφάσεων, να αντιληφθούν την πληροφορία αποτελεσματικότερα. Εργαλεία χαρτογραφικής οπτικοποίησης μπορούν να αποδώσουν γραφικές αναπαραστάσεις χωρικών δεδομένων, όπως η παρακολούθηση της κυκλοφορίας και η διαχείριση αστικών υποδομών, καθιστώντας τα δεδομένα πιο προσβάσιμα και εύληπτα. Επιπλέον, η προχωρημένη ανάλυση και εξαγωγή πληροφοριών από δεδομένα AI επιτρέπει τη δημιουργία ειδοποιήσεων γεωγραφικού φράχτη σε πραγματικό χρόνο. Αυτό σημαίνει ότι το σύστημα μπορεί να εντοπίζει και να αντιδρά άμεσα σε συγκεκριμένα γεγονότα ή συνθήκες, παρέχοντας έγκαιρες προειδοποιήσεις και ειδοποιήσεις. Η ολοκληρωμένη αξιοποίηση ΓΠΣ και AI μπορεί να φέρει σημαντικά οφέλη στον τομέα της ανάλυσης δεδομένων και της υποστήριξης αποφάσεων, ενώ διαθέτει ένα ευρύ φάσμα εφαρμογών σε διάφορους κλάδους, όπως οι μεταφορές, η διαχείριση πόλεων, η υγεία, η περιβαλλοντική παρακολούθηση και πολλοί άλλοι. (Vozenilek, Vit. 2009)

7 Περιπτώσεις χρήσης

Στην παρούσα ενότητα καταγράφονται ενδεικτικές περιπτώσεις χρήσης του συνδυασμού ΓΠΣ και τεχνητής νοημοσύνης.

7.1 Παγκόσμια Προγράμματα Χαρτογράφησης Καλύψεων και Χρήσεων Γης

Η ανάλυση τηλεπισκοπικών δεδομένων αποτελεί βασικό εργαλείο για τη χαρτογράφηση των καλύψεων και χρήσεων γης σε παγκόσμια κλίμακα, παρέχοντας ζωτικής σημασίας πληροφορίες για την παρακολούθηση των μεταβολών του περιβάλλοντος και τη βιώσιμη διαχείριση των φυσικών πόρων. Οι τηλεπισκοπικές τεχνικές επιτρέπουν τη δημιουργία λεπτομερών χαρτών που αποτυπώνουν τις αλλαγές στη χρήση γης, όπως η αποψίλωση των δασών, η αστική εξάπλωση και οι γεωργικές δραστηριότητες, προσφέροντας μια ολοκληρωμένη εικόνα για τις χωρικές και χρονικές μεταβολές των τοπίων. Επιπλέον, οι χάρτες αυτοί μπορούν να υποστηρίξουν τη χάραξη περιβαλλοντικών πολιτικών, τη διαχείριση των υδάτινων πόρων και την παρακολούθηση της βιοποικιλότητας, συμβάλλοντας στην επίτευξη των στόχων της βιώσιμης ανάπτυξης. Στην ερευνητική εργασία του Δουνδουλάκη (2020), εξετάστηκαν παγκόσμια προγράμματα χαρτογράφησης καλύψεων και χρήσεων γης, όπως το Global Land Cover 2000 (GLC2000), το GlobCover και το GlobeLand30. Αυτά τα προγράμματα αξιοποίησαν δορυφορικές εικόνες από αισθητήρες όπως ο MERIS και ο Landsat, εφαρμόζοντας προηγμένες τεχνικές ταξινόμησης εικόνων και εκτίμησης ακρίβειας. Το GLC2000 ήταν ένα από τα πρώτα παγκόσμια προγράμματα χαρτογράφησης που βασίστηκαν σε δεδομένα από τον αισθητήρα VEGETATION του δορυφόρου SPOT-4, ενώ το GlobCover ανέπτυξε πιο λεπτομερείς χάρτες, χρησιμοποιώντας δεδομένα από τον αισθητήρα MERIS του Envisat. Το πιο πρόσφατο πρόγραμμα, GlobeLand30, εισήγαγε σημαντικές βελτιώσεις στην ανάλυση, αξιοποιώντας δεδομένα από τους δορυφόρους Landsat και HJ-1 για τη δημιουργία χαρτών με χωρική ανάλυση 30 μέτρων, επιτρέποντας πιο ακριβείς εκτιμήσεις για τη χρήση και κάλυψη της γης. Η τηλεπισκόπηση επιτρέπει τη συνεχή ενημέρωση αυτών των χαρτών, διευκολύνοντας την παρακολούθηση των φυσικών και ανθρωπογενών αλλαγών σε μεγάλη κλίμακα. Η ικανότητα των δορυφορικών συστημάτων να παρέχουν δεδομένα σε τακτά χρονικά διαστήματα συμβάλλει στη διαχείριση των φυσικών πόρων, στον χωροταξικό σχεδιασμό και στην ανάλυση των επιπτώσεων της κλιματικής αλλαγής. Για παράδειγμα, οι δορυφορικές εικόνες μπορούν να χρησιμοποιηθούν για την ανίχνευση παράνομων αποψιλώσεων, την παρακολούθηση των πλημμυρών και την εκτίμηση της διάβρωσης του εδάφους, παρέχοντας στους υπεύθυνους λήψης αποφάσεων κρίσιμες πληροφορίες για την προστασία του περιβάλλοντος. Η έρευνα ανέδειξε τις διαφορές μεταξύ των προγραμμάτων χαρτογράφησης ως προς τη μεθοδολογία, την ακρίβεια και την ανάλυση των δεδομένων που χρησιμοποιούν. Συγκεκριμένα, το GlobeLand30 υπερέχει λόγω της υψηλής χωρικής ανάλυσής του, επιτρέποντας πιο ακριβείς ταξινομήσεις των καλύψεων γης. Ενώ το GLC2000 και το GlobCover παρέχουν χρήσιμα δεδομένα σε παγκόσμια κλίμακα, η χαμηλότερη χωρική τους ανάλυση περιορίζει την ικανότητά τους να ανιχνεύουν λεπτομερείς μεταβολές στο τοπίο. Ως αποτέλεσμα, το GlobeLand30 αποτελεί σήμερα ένα από τα πιο προηγμένα εργαλεία για τη χαρτογράφηση της γης, προσφέροντας δεδομένα υψηλής ακρίβειας που μπορούν να χρησιμοποιηθούν σε ένα ευρύ φάσμα εφαρμογών, από την περιβαλλοντική διαχείριση έως τον αστικό σχεδιασμό και τη γεωργική παρακολούθηση (Δουνδουλάκης, 2020).

7.2 Γεωργία Ακριβείας και Μηχανική Μάθηση

Η εφαρμογή τεχνικών μηχανικής μάθησης και εξόρυξης δεδομένων στη γεωργία ακριβείας αποτελεί ένα σημαντικό πεδίο που επιτρέπει τη βελτίωση της παραγωγής μέσω τηλεπισκοπικών δεδομένων, προσφέροντας στους γεωργούς και τους ερευνητές πολύτιμες πληροφορίες για τη διαχείριση των καλλιεργειών. Η χρήση αισθητήρων απομακρυσμένης παρατήρησης, συνδυασμένων με εξελιγμένους αλγόριθμους μηχανικής μάθησης, συμβάλλει στην ακριβέστερη εκτίμηση παραμέτρων όπως η υγεία της βλάστησης, η κατανομή των καλλιεργειών και η παρουσία αγριόχορτων ή εντόμων που μπορούν να επηρεάσουν την παραγωγή. Στη μελέτη της Δημαράκη (2022), χρησιμοποιήθηκαν δεδομένα από μη επανδρωμένα ιπτάμενα οχήματα (UAVs) για την ανάλυση της καλλιέργειας κρόκου στην περιοχή της Κοζάνης, η οποία αποτελεί μια από τις σημαντικότερες ζώνες παραγωγής κρόκου στην Ελλάδα. Τα UAVs εξοπλίστηκαν με πολυφασματικές κάμερες που επέτρεψαν τη συλλογή δεδομένων υψηλής ανάλυσης, τα οποία στη συνέχεια επεξεργάστηκαν με τεχνικές φωτογραμμετρίας και ανάλυσης εικόνας. Μέσω της εφαρμογής υπολογιστικών μεθόδων, όπως η εξαγωγή του δείκτη βλάστησης NDVI (Normalized Difference Vegetation Index), ήταν δυνατή η δημιουργία λεπτομερών χαρτών που απεικόνιζαν τη γενική κατάσταση της καλλιέργειας. Οι εικόνες που συλλέχθηκαν κατηγοριοποιήθηκαν με δύο διαφορετικές προσεγγίσεις: την ανάλυση ανά εικονοστοιχείο και την αντικειμενοστραφή ανάλυση εικόνας (OBIA). Η πρώτη μέθοδος εστίασε στην επεξεργασία των δεδομένων σε επίπεδο μεμονωμένων pixel, παρέχοντας μεγαλύτερη λεπτομέρεια στη χαρτογράφηση της βλάστησης, ενώ η δεύτερη βασίστηκε στην ομαδοποίηση γειτονικών pixel για τη δημιουργία πιο ομοιογενών κατηγοριών. Για την εξαγωγή συμπερασμάτων, εκπαιδεύτηκαν δύο αλγόριθμοι μηχανικής μάθησης: τα Τυχαία Δάση (Random Forests) και τα Νευρωνικά Δίκτυα (Neural Networks). Τα αποτελέσματα της έρευνας έδειξαν ότι η μέθοδος ανάλυσης ανά εικονοστοιχείο παρουσίασε υψηλότερη ακρίβεια, καθώς κατάφερε να εντοπίσει με μεγαλύτερη ευαισθησία τις διαφορές μεταξύ των φυτών και των υπολοίπων στοιχείων του εδάφους. Τα Τυχαία Δάση αναδείχθηκαν ως η πιο αποδοτική τεχνική, καταγράφοντας ποσοστά ακρίβειας μεταξύ 88% και 100%, ενώ τα Νευρωνικά Δίκτυα παρουσίασαν χαμηλότερες επιδόσεις, ιδιαίτερα στην αντικειμενοστραφή ανάλυση, με ποσοστά που κυμάνθηκαν από 81% έως 87%. Συμπερασματικά, η χρήση τηλεπισκοπικών δεδομένων σε συνδυασμό με τεχνικές μηχανικής μάθησης μπορεί να αποτελέσει ένα ισχυρό εργαλείο για την παρακολούθηση των γεωργικών εκτάσεων, επιτρέποντας τη λήψη πιο στοχευμένων αποφάσεων για τη βελτίωση της παραγωγής (Δημαράκη, 2022).

7.3 Επιχειρηματική Νοημοσύνη μέσω Τηλεπισκόπησης και Μηχανικής Μάθησης

Η χρήση δορυφορικών δεδομένων σε συνδυασμό με τεχνικές μηχανικής μάθησης μπορεί να αποτελέσει ένα ισχυρό εργαλείο για την υποστήριξη επιχειρηματικών αποφάσεων, ιδίως στον τομέα της ανάλυσης του αστικού περιβάλλοντος. Τα σύγχρονα δορυφορικά δεδομένα παρέχουν τεράστιες ποσότητες πληροφορίας που, με την κατάλληλη επεξεργασία, μπορούν να προσφέρουν πολύτιμες γνώσεις σχετικά με τις χωρικές και κοινωνικοοικονομικές αλλαγές στις πόλεις. Η ανάλυση αυτών των δεδομένων μπορεί να συμβάλει στην αναγνώριση τάσεων, στη βελτίωση της πολεοδομικής ανάπτυξης και στην εκτίμηση του οικονομικού δυναμικού συγκεκριμένων περιοχών. Επιπλέον, η εφαρμογή τεχνολογιών τεχνητής νοημοσύνης και μηχανικής μάθησης επιτρέπει την αυτοματοποίηση της διαδικασίας ανάλυσης, μειώνοντας το ανθρώπινο σφάλμα και βελτιώνοντας την ακρίβεια των

προβλέψεων. Στη διπλωματική εργασία του Κυριάκου (2022), προτάθηκε ένα καινοτόμο ανοιχτού κώδικα λογισμικό που επιτρέπει τον εντοπισμό εγκαταλελειμμένων κτιρίων μέσω της ανάλυσης δορυφορικών εικόνων. Η μέθοδος που εφαρμόστηκε βασίζεται στη χρήση εξειδικευμένων δεικτών, όπως τα Nighttime Lights και Built-Up Area Indices, οι οποίοι παρέχουν σημαντικές πληροφορίες σχετικά με τη φωτεινότητα και τη δόμηση μιας περιοχής αντίστοιχα. Οι δείκτες αυτοί χρησιμοποιούνται για να διακρίνουν περιοχές με ενεργή οικονομική και κοινωνική δραστηριότητα από εκείνες που παρουσιάζουν μειωμένη χρήση ή εγκατάλειψη. Το λογισμικό αναλύει τα δεδομένα αυτά και εντοπίζει κτίρια που παρουσιάζουν σημάδια εγκατάλειψης, κάτι που μπορεί να έχει σημαντικές επιπτώσεις στον αστικό σχεδιασμό και στις στρατηγικές αναζωογόνησης του αστικού χώρου. Η προσέγγιση αυτή έχει ιδιαίτερη σημασία για μικρομεσαίες επιχειρήσεις και υπεύθυνους χάραξης πολιτικής, καθώς διευκολύνει τη λήψη αποφάσεων σχετικά με την αξιοποίηση ανεκμετάλλευτων ακινήτων και τη στοχευμένη επένδυση σε περιοχές που χρειάζονται αναβάθμιση. Για παράδειγμα, οι επιχειρήσεις μπορούν να χρησιμοποιήσουν τα δεδομένα αυτά για να εντοπίσουν ευκαιρίες ανάπτυξης σε υποβαθμισμένες περιοχές, ενώ οι τοπικές αρχές μπορούν να σχεδιάσουν προγράμματα αποκατάστασης κτιρίων ή να παρέχουν φορολογικά κίνητρα για την επανεκμετάλλευση εγκαταλελειμμένων χώρων. Η μελέτη ανέδειξε τη σημασία της τεχνητής νοημοσύνης και των εργαλείων τηλεπισκόπησης στη διαχείριση των "Big Satellite Data", προσφέροντας βελτιωμένες δυνατότητες ανάλυσης και πρόβλεψης. Τα πειράματα που διεξήχθησαν έδειξαν ότι η προτεινόμενη προσέγγιση είναι πολλά υποσχόμενη, με υψηλά ποσοστά επιτυχίας στον εντοπισμό εγκαταλελειμμένων κτιρίων. Επιπλέον, η χρήση ανοικτών δεδομένων και η προσβασιμότητα του λογισμικού σε ερευνητές και επαγγελματίες καθιστά τη μέθοδο αυτή ιδιαίτερα ελκυστική για περαιτέρω ανάπτυξη και εφαρμογή σε διαφορετικές περιοχές και συνθήκες. Συμπερασματικά, η συνδυαστική αξιοποίηση της τηλεπισκόπησης και της μηχανικής μάθησης μπορεί να συμβάλει ουσιαστικά στη διαμόρφωση βιώσιμων και ευφών στρατηγικών για την αναζωογόνηση των πόλεων και τη βελτίωση της ποιότητας ζωής (Κυριάκος, 2022).

7.4 Ανίχνευση Οροφών για Φωτοβολταϊκά Συστήματα μέσω Μηχανικής Μάθησης

Η αξιοποίηση των τεχνικών τηλεπισκόπησης και μηχανικής μάθησης για την ανίχνευση οροφών σε αστικά περιβάλλοντα αποτελεί κρίσιμη εφαρμογή για τον ενεργειακό σχεδιασμό και την προώθηση βιώσιμων λύσεων για την παραγωγή ηλεκτρικής ενέργειας. Ο εντοπισμός των διαθέσιμων οροφών και η αξιολόγηση της καταλληλότητάς τους για την εγκατάσταση φωτοβολταϊκών συστημάτων συμβάλλουν στη βέλτιστη αξιοποίηση της ηλιακής ενέργειας και στη μείωση της εξάρτησης από συμβατικές πηγές ενέργειας. Με τη χρήση δορυφορικών εικόνων υψηλής ανάλυσης και τεχνικών βαθιάς μάθησης, καθίσταται δυνατός ο αυτόματος εντοπισμός και η ταξινόμηση των οροφών, επιτρέποντας τη χαρτογράφηση περιοχών με υψηλό δυναμικό για την παραγωγή ηλιακής ενέργειας. Στη μελέτη της Βαλδεσέρα (2023), αναπτύχθηκε ένα μοντέλο βαθιάς μάθησης (Deep Learning) με στόχο την αυτόματη εξαγωγή οροφών κτιρίων από δορυφορικές εικόνες και την αξιολόγηση της δυνατότητας εγκατάστασης φωτοβολταϊκών πάνελ. Ο αλγόριθμος που αναπτύχθηκε βασίστηκε σε Συνελικτικά Νευρωνικά Δίκτυα (CNNs) και μεθόδους υπολογιστικής όρασης, επιτρέποντας την ανίχνευση και ταξινόμηση των οροφών με υψηλή ακρίβεια. Τα δεδομένα που χρησιμοποιήθηκαν προήλθαν από δορυφορικές εικόνες υψηλής χωρικής ανάλυσης, επιτρέποντας τη λεπτομερή χαρτογράφηση της αστικής περιοχής. Η μεθοδολογία

περιλάμβανε προεπεξεργασία των εικόνων, τμηματοποίηση των οροφών και τελική κατηγοριοποίηση, με στόχο τη διάκριση των επίπεδων από τις επικλινείς επιφάνειες, καθώς και την αναγνώριση εμποδίων όπως καμινάδες, κεραίες και άλλες δομές που μπορεί να επηρεάζουν την απόδοση των φωτοβολταϊκών συστημάτων. Η μελέτη επικεντρώθηκε στην περιοχή της Νίκαιας στην Αττική, όπου το αστικό τοπίο χαρακτηρίζεται από μεγάλη ποικιλομορφία στις κατασκευές των κτιρίων και την ύπαρξη πολλών διαθέσιμων οροφών που μπορούν να αξιοποιηθούν για ενεργειακούς σκοπούς. Μέσω της ανάλυσης των γεωμετρικών χαρακτηριστικών των οροφών και της εκτίμησης της ηλιακής ακτινοβολίας που δέχονται, η προσέγγιση αυτή μπορεί να υποστηρίξει στρατηγικές αποφάσεις για την τοποθέτηση φωτοβολταϊκών συστημάτων σε μεγάλη κλίμακα. Επιπλέον, η τεχνική αυτή μπορεί να επεκταθεί για την εκτίμηση της ενεργειακής απόδοσης και την πρόβλεψη του δυναμικού παραγωγής ηλεκτρικής ενέργειας, ενσωματώνοντας δεδομένα όπως οι καιρικές συνθήκες και η σκίαση από γειτονικά κτίρια ή φυσικά εμπόδια. Τα αποτελέσματα της έρευνας έδειξαν ότι η μέθοδος μπορεί να εφαρμοστεί σε διάφορες κλίμακες, από μεμονωμένα κτίρια μέχρι ολόκληρες πόλεις, καθιστώντας την ένα ισχυρό εργαλείο για τον σχεδιασμό πολιτικών προώθησης ανανεώσιμων πηγών ενέργειας. Ο συνδυασμός τηλεπισκόπησης και μηχανικής μάθησης προσφέρει μια αυτοματοποιημένη και αποδοτική λύση για την ανίχνευση και αξιολόγηση των ταρατσών, μειώνοντας την ανάγκη για επίγειες επιθεωρήσεις και αυξάνοντας την ακρίβεια των αποτελεσμάτων. Η συγκεκριμένη μεθοδολογία μπορεί να ενισχύσει τις πρωτοβουλίες για τη βιώσιμη αστική ανάπτυξη και τη μετάβαση προς ένα πιο πράσινο ενεργειακό μοντέλο, ενσωματώνοντας την τεχνητή νοημοσύνη στην ανάλυση και λήψη αποφάσεων για την εκμετάλλευση της ηλιακής ενέργειας (Βαλδεσέρα, 2023).

8 Μελέτη περίπτωσης

Στο πλαίσιο της παρούσας διπλωματικής εργασίας εκπονήθηκε μία μελέτη περίπτωσης με εφαρμογή του συνδυασμού ΓΠΣ και μεθόδων τεχνητής νοημοσύνης. Για τη μελέτη χρησιμοποιήθηκαν δεδομένα από μία περιοχή της Κορινθίας. Η συγκεκριμένη περιοχή επιλέχθηκε αφ' ενός λόγω εντοπιότητας του συγγραφέα, δευτερευόντως λόγω της ποικιλομορφίας που εμφανίζει η περιοχή σε χρήσεις γης, καθώς και σε φυσικά και αστικά χαρακτηριστικά που αξιοποιούνται στο πλαίσιο της μηχανικής μάθησης.

8.1 Γεωγραφικά στοιχεία

Ο νομός Κορινθίας έχει συνολική έκταση 2.290 km² και καταλαμβάνει το βορειοανατολικό άκρο της Πελοποννήσου και μικρό τμήμα της Στερεάς Ελλάδας (Γεράνεια όρη). Πρωτεύουσά του είναι η Κόρινθος. Βόρεια βρέχεται σ' όλο το μήκος του από τον Κορινθιακό κόλπο και ανατολικά από το Σαρωνικό, οι οποίοι επικοινωνούν μεταξύ τους με την τεχνητή διώρυγα της Κορίνθου (Ισθμός Κορίνθου). Νότια συνορεύει με το νομό Αργολίδας και Αρκαδίας και δυτικά με το νομό Αχαΐας. Το τμήμα του νομού που βρίσκεται στη Στερεά Ελλάδα, συνορεύει με το νομό Αττικής.

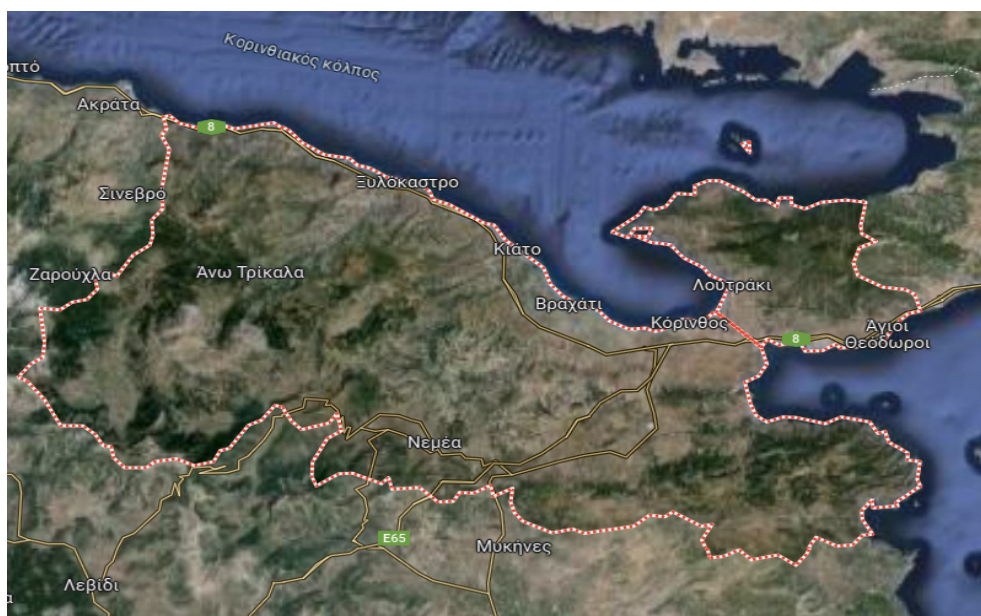


Εικόνα 15: Γεωγραφική θέση του νομού Κορινθίας (πηγή elwikipedia.org)

8.1.1 Τοπογραφικά Στοιχεία

Το μεγαλύτερο μέρος της Κορινθίας είναι ορεινό - ημιορεινό και μόνο το βόρειο τμήμα της είναι πεδινό. Στο νότιο τμήμα του νομού αναπτύσσονται εσωτερικές πεδιάδες και οροπέδια. Στο νομό περιλαμβάνονται μια σειρά από όρη, όπως η Κυλλήνη (Ζήρια), με ύψος 2.376 μ., τα Γεράνεια όρη με ύψος 1.351μ., η Τραπεζώνα με ύψος 1.134 μ., το όρος Τρίκορφο με ύψος 801 μ., το Τρωτό όρος ύψους 722 μ., το Φαρμακά ύψους 1.615 μ., ο Ολίγυρτος ύψους 1.935 μ., το Γερόντιο όρος ύψους 1.756 μ., η Όρυξις ύψους 1.814 μ. και το όρος Ντουρντουβάνια ύψους 2019 μ. Στο εσωτερικό των οροσειρών αυτών περικλείονται μία σειρά από οροπέδια, τα οποία από Ανατολάς προς Δυσμάς είναι αυτά του Σοφικού, του Αγίου-Ιωάννη, της Αγίας Τριάδας, του Αγιονορίου, της Σκοτεινής, της Στυμφαλίας, της Σκαφιδιάς και του Φενεού. Στο μέσον του νομού της Κορινθίας βρίσκονται οι παρακάτω λόφοι και τα όρη, τα οποία από τα ανατολικά προς τα δυτικά είναι τα ακόλουθα: Όνεια όρη (583 μ.), Ακροκόρινθος (574 μ.), Σκιώνα

(764 μ.), Φωκάς (872 μ.), Τρικάρανο (806 μ.), Γαβριάς (1.207 μ.), Βέσιζα (1.206 μ.), Μικρό όρος (1.600 μ.), Ευρωσίνη (1.473 μ.). Μαύρο όρος (1.746 μ.). Στα όρη αυτά της ορεινής και λοφώδους Κορινθίας εντοπίζονται μικρές πεδιάδες όπως του Γαλατακίου, των Αθικίων, του Χιλιομοδίου, του Κοντόσταυλου, της Αρχαίας Νεμέας, της Νεμέας και του Ασπρόκαμπου. Από την περιοχή του χωριού Σούλι και δυτικά δεν έχουμε αντίστοιχα μικρές πεδιάδες που να παρεμβάλλονται ανάμεσα σε όρη, αλλά επικρατούν απότομες κοιλάδες. Στο βόρειο τμήμα της Κορινθίας αναπτύσσεται η παραλιακή πεδιάδα της Βόχας, που στα ανατολικά της συνδέεται με τις πεδιάδες της παλαιάς Κορίνθου και των Εξαμιλίων- Κεγχρεών.



Εικόνα 16: Γεωφυσική απεικόνιση του νομού Κορινθίας (πηγή Google Maps)

8.1.2 Δημογραφικά Στοιχεία

Σύμφωνα με το Πρόγραμμα «Καλλικράτης», ο νομός Κορινθίας αποτελεί την Περιφερειακή Ενότητα (Π.Ε.) Κορινθίας (με έδρα την Κόρινθο) της Περιφέρειας Πελοποννήσου της Αποκεντρωμένης Διοίκησης Πελοποννήσου, Δυτικής Ελλάδας και Ιονίου. Η Π.Ε. αποτελείται από 6 Δήμους:

- Δήμος Κορινθίων
- Δήμος Σικυνίων
- Δήμος Λουτρακίου - Περαιώς - Αγίων Θεοδώρων
- Δήμος Βέλου – Βόχας
- Δήμος Νεμέας
- Δήμος Ξυλοκάστρου – Ευρωστίνης

Ο νομός έχει πληθυσμό 138.144 κατοίκους, σύμφωνα την απογραφή του 2021

8.1.3 Υδρολογικά Στοιχεία-Κλίμα

Η Κορινθία έχει μεγάλο αριθμό χειμάρρων οι οποίοι δεν συνενώνονται σε μεγάλους ποταμούς. Η μοναδική λίμνη του νομού είναι η λίμνη της Στυμφαλίας. Στο δυτικό άκρο της απόληξης των Γερανείων, βρίσκεται η λίμνη Βουλιαγμένης ή Ηραίου, η οποία συνδέεται μέσω ενός μικρού ανοίγματος με τη θάλασσα. Το κλίμα του νομού είναι ξηρό, με ήπιους χειμώνες και δροσερά καλοκαίρια. Οι βροχοπτώσεις

είναι μεγαλύτερες στο ορεινό δυτικό τμήμα (600 χιλιοστά κατ' έτος), ελαττώνονται όμως στο ανατολικό (Κόρινθος, 413 χιλιοστά (meteo)).

8.1.4 Οικονομικές Δραστηριότητες

Η Κόρινθος και το Λουτράκι είναι μεγάλος βιομηχανικός κόμβος σε εθνικό επίπεδο. Το διυλιστήριο της Κορίνθου είναι ένα από τα μεγαλύτερα συγκροτήματα διύλισης πετρελαίου στην Ευρώπη. Επίσης παράγονται κεραμικά πλακίδια, καλώδια χαλκού, γύψος, δερμάτινα προϊόντα, μάρμαρο, προϊόντα κρέατος, ιατρικός εξοπλισμός, μεταλλικό νερό, ποτά και αγροτικά προϊόντα. Οι κάτοικοί της επίσης ασχολούνται με τη γεωργία την κτηνοτροφία, αλλά και με τον τουρισμό.



Εικόνα 17: Χαρτοσύνθεση περιοχής μελέτης με όμορους Δήμους

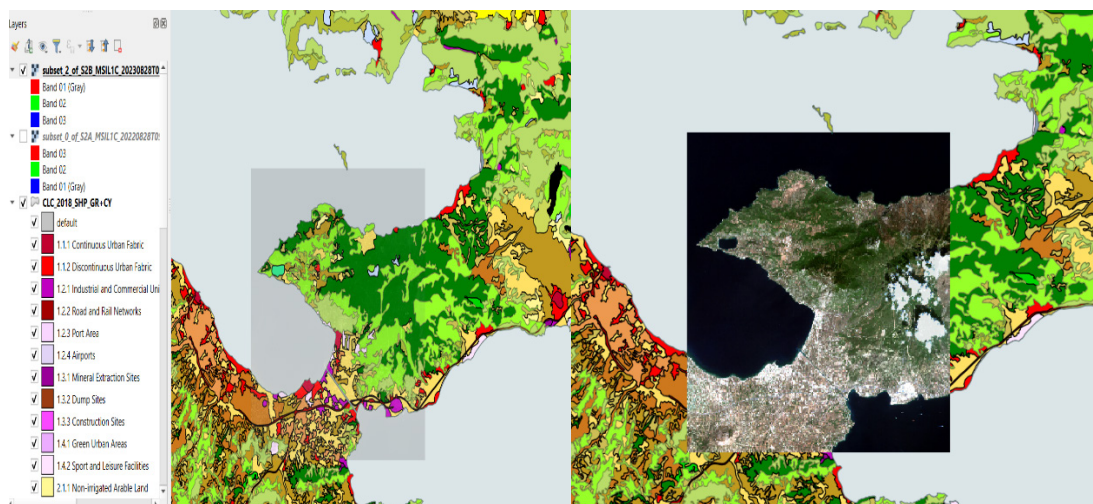
8.2 Χρήσεις Γης

Για την εύρεση των χρήσεων γης της περιοχής, χρησιμοποιήθηκε το θεματικό επίπεδο Corine Land Cover (CLC) του 2018¹². Η χρήση των δεδομένων του CORINE Land Cover πριν από τη μηχανική μάθηση είναι κρίσιμη, καθώς παρέχει μια σαφή εικόνα των χρήσεων γης, βελτιώνει την ποιότητα των δεδομένων εκπαίδευσης και αυξάνει την ακρίβεια του μοντέλου. Ο χάρτης αυτός των χρήσεων γης αναλύθηκε με χρήση του ΣΓΠ ανοικτού κώδικα QGIS. Για να γίνει η εύρεση των χρήσεων γης εκτελέστηκε το εργαλείο *Select Features by Polygon* στο παραπάνω θεματικό επίπεδο και ως πολύγωνο σχεδιάστηκε το περίγραμμα της δορυφορικής εικόνας. Τα αποτελέσματα της εκτέλεσης του παραπάνω εργαλείου, αναπαρίστανται στον παρακάτω πίνακα.

¹² πηγή: <https://land.copernicus.eu>

Κωδικός	Περιγραφή
1.1.1	Συνεχής αστικός ιστός
1.1.2	Ασυνεχής αστικός ιστός
1.2.1	Βιομηχανικές και εμπορικές ζώνες
1.2.2	Οδικά και σιδηροδρομικά δίκτυα
1.4.2	Εγκαταστάσεις αθλητισμού και αναψυχής
2.1.1	Μη αρδευόμενη αρώσιμη γη
2.2.1	Αμπελώνες
2.2.2	Οπωροφόρα δένδρα και φυτείες με σαρκώδεις καρπούς
2.2.3	Ελαιώνες
2.3.1	Λιβάδια
2.4.2	Σύνθετες καλλιέργειες
2.4.3	Γη που χρησιμοποιείται κυρίως για γεωργία μαζί με σημαντικά τμήματα φυσικής βλάστησης
3.1.2	Δάσος κωνοφόρων
3.2.1	Φυσικοί βοσκότοποι
3.2.3	Σκληροφυλλική βλάστηση
3.2.4	Μεταβατικές δασώδεις και θαμνώδεις εκτάσεις
3.3.3	Εκτάσεις με αραιή βλάστηση
5.2.1	Παράκτιες Λιμνοθάλασσες
5.2.3	Θάλασσες και ωκεανοί

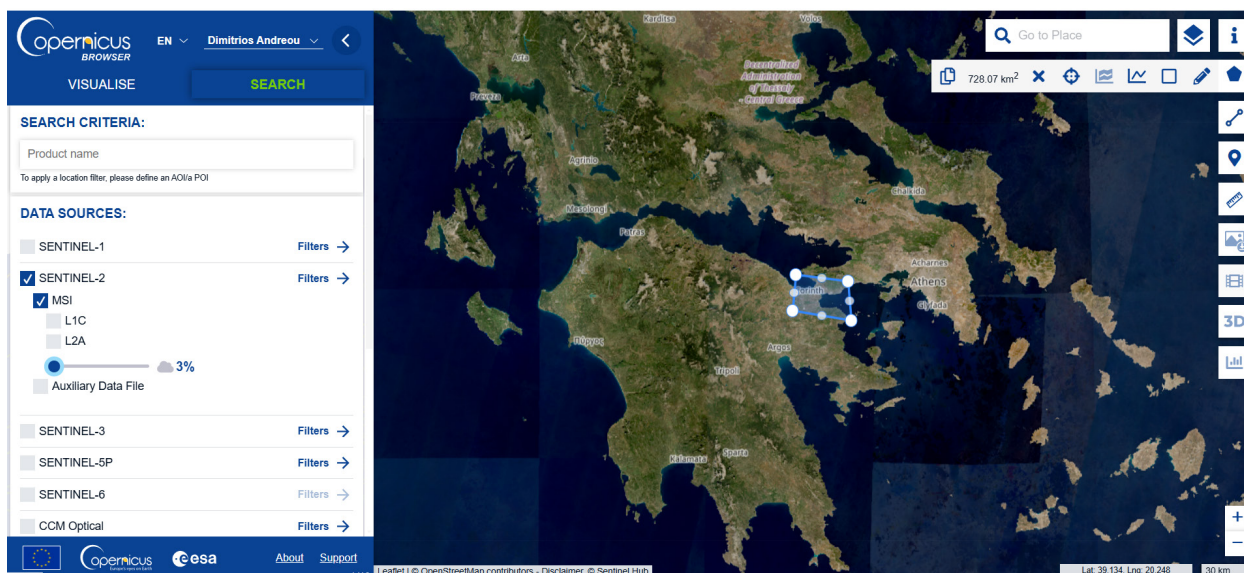
Πίνακας 2: Κατηγορίες χρήσεων γης CLC 2018 που υπάρχουν στην περιοχή μελέτης.



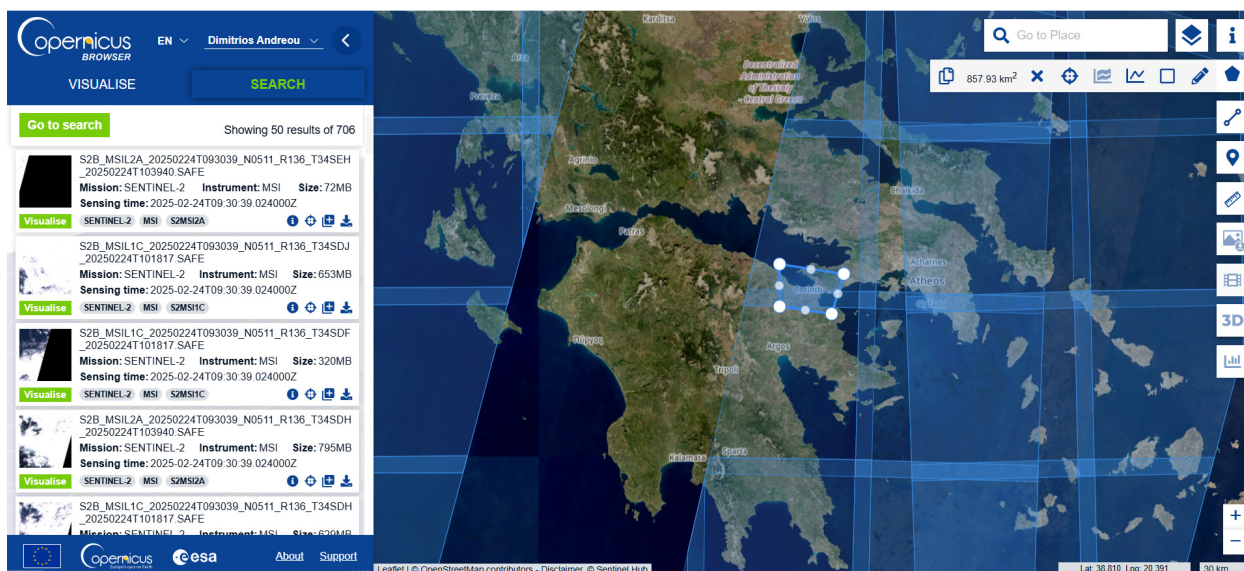
Εικόνα 18: Χαρτοσύνθεση με δορυφορική εικόνα και χρήσεις γης CLC 2018.

8.3 Λήψη εικόνων

Στη συνέχεια πραγματοποιήθηκε εξέταση και λήψη δορυφορικών εικόνων που αφορούν την ευρύτερη περιοχή-στόχο, από το σύνολο δεδομένων SENTINEL 2, μέσω της πλατφόρμας «Copernicus Open Access Hub». Η διαδικασία φαίνεται στην Εικόνα 19 και την Εικόνα 20



Εικόνα 19: Επιλογή πολυγώνου επιθυμητής περιοχής και δορυφόρου



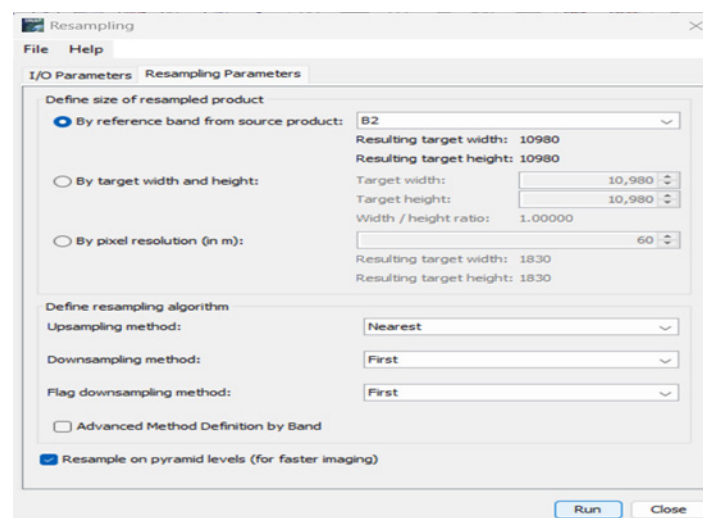
Εικόνα 20: Διαθέσιμες δορυφορικές φωτογραφίες

Τελικά επιλέχθηκε μια φωτογραφία από τις 28 Αυγούστου του 2023 η οποία διέθετε όλα τα επιθυμητά χαρακτηριστικά, δηλαδή να καλύπτει πολλές χρήσεις γης και να χαρακτηρίζεται από χαμηλή νεφοκάλυψη.

8.4 Επαναδειγματοληψία και αποκοπή εικόνας

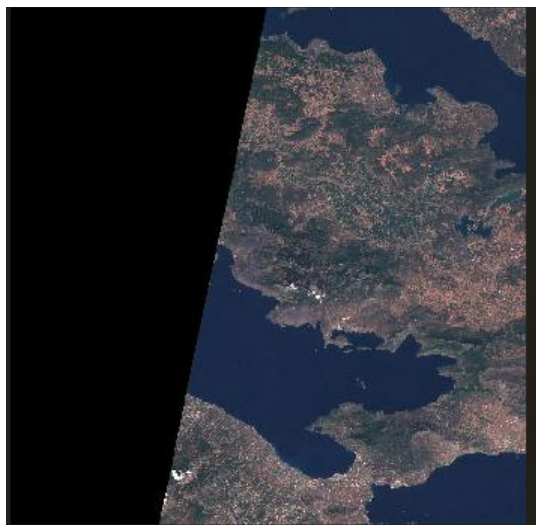
Ο Sentinel-2 καταγράφει δεδομένα σε 13 φασματικές ζώνες, αλλά οι φασματικές ζώνες αυτές έχουν διαφορετικές χωρικές αναλύσεις (10m, 20m, 60m). Προκειμένου να είναι δυνατή η συνδυαστική ανάλυσή τους, είναι απαραίτητο να εκτελεστεί η διαδικασία της επαναδειγματοληψίας. Η επαναδειγματοληψία (resampling) είναι η διαδικασία αλλαγής της χωρικής ανάλυσης ενός δεδομένου (εικόνας ή καναλιού) ώστε να αποκτήσει νέο μέγεθος pixel. Στην περίπτωση του Sentinel-2, επειδή τα κανάλια καταγράφονται σε διαφορετικές αναλύσεις (10m, 20m, 60m), η επαναδειγματοληψία είναι

απαραίτητη για την ομοιόμορφη ανάλυση όλων των καναλιών, ώστε να μπορούν να συνδυαστούν σε κοινές αναλύσεις και να επιτρέψουν αξιόπιστους υπολογισμούς και αναλύσεις (π.χ. δεικτών βλάστησης). Για την υλοποίηση της επιλέχθηκε το λογισμικό SNAP. Το SNAP (Sentinel Application Platform) είναι ένα ελεύθερο λογισμικό επεξεργασίας δορυφορικών δεδομένων που αναπτύχθηκε από τον Ευρωπαϊκό Οργανισμό Διαστήματος (ESA). Παρέχει εξειδικευμένα εργαλεία για την επεξεργασία προϊόντων Sentinel, συμπεριλαμβανομένης της διόρθωσης ατμόσφαιρας, της γεωμετρικής διόρθωσης και της επαναδειγματοληψίας. Επιλέχθηκε λόγω της πλήρους συμβατότητάς του με τα δεδομένα Sentinel, της αξιοπιστίας του και της δωρεάν διάθεσής του. Επιλέχθηκε το κανάλι B2 με χωρική ανάλυση στα 10μ για να γίνει αναδόμηση. Το κανάλι B2 (μπλε) έχει χωρική ανάλυση 10 μέτρων και αποτελεί ένα από τα κανάλια υψηλότερης ανάλυσης του Sentinel-2. Επιλέχθηκε ως βάση για την επαναδειγματοληψία ώστε όλα τα υπόλοιπα κανάλια να προσαρμοστούν σε αυτήν την ανάλυση, επιτυγχάνοντας καλύτερη ευκρίνεια, λεπτομέρεια και ακρίβεια στις επόμενες αναλύσεις και συνδυασμούς καναλιών. Οι παράμετροι της επαναδειγματοληψίας παρουσιάζονται στην Εικόνα 21.



Εικόνα 21: Αναδόμηση στο κανάλι B2

Παρακάτω θα παρουσιαστεί η φωτογραφία που λήφθηκε από το Copernicus με τα χαρακτηριστικά της.

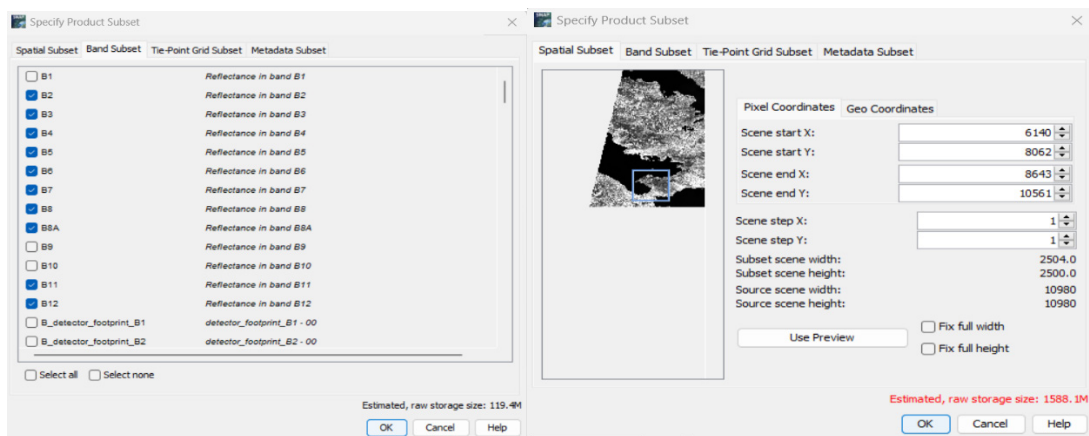


Εικόνα 22: Αρχική Δορυφορική Εικόνα

Όνομα αρχικής δορυφορικής εικόνας	S2B_MSIL1C_20230828T090559_N0509_R050_T34SFH_20230828T111807
Ημερομηνία/ Ώρα λήψης	2023-08-28/09:05:59.024ZONE
Γραμμές	10980
Pixels ανά γραμμή	10980
Συντεταγμένες περιοχής(EPSC:3263 4 - WGS 84 / UTM zone 34N)	Upper Left (22° 50' 28.22"E, 38° 6' 27.54"N) Lower Left (22° 50' 7.19"E, 37° 52' 58.11"N) Upper Right (23° 07' 34.57"E, 38° 6' 10.20"N) Lower Right (23° 07' 11.64"E, 37° 52' 40.89"N) Center (22° 58' 50.48"E, 37°59'33.69"N)

Όπως φαίνεται και από την πλατφόρμα της υπηρεσίας Copernicus, οι διαστάσεις της εικόνας είναι πολύ μεγάλες συνεπώς επιλέχθηκε μια περιοχή μελέτης και αποκόπηκε. Για τη μελέτη των χρήσεων γης στην περιοχή ενδιαφέροντος, επιλέχθηκαν δέκα συγκεκριμένα κανάλια του δορυφόρου Sentinel-2: το μπλε (B2), το πράσινο (B3), το κόκκινο (B4), τα τρία vegetation red edge (B5, B6, B7), το ευρύ εγγύς

υπέρυθρο (B8), το στενό εγγύς υπέρυθρο (B8A), καθώς και δύο κανάλια στο βραχύ υπέρυθρο (SWIR, B11 και B12). Η επιλογή αυτών των καναλιών βασίστηκε στην ανάγκη καταγραφής κρίσιμων χαρακτηριστικών του εδάφους και της βλάστησης, τα οποία είναι απαραίτητα για τη σαφή διάκριση των διαφορετικών τύπων κάλυψης και χρήσης γης. Τα κανάλια B2, B3 και B4 (μπλε, πράσινο και κόκκινο) είναι τα βασικά κανάλια του ορατού φάσματος και επιτρέπουν τόσο τη δημιουργία ψευδοχρωματικών και αληθοφανών συνθέσεων όσο και την εφαρμογή βασικών αναλύσεων. Τα κανάλια B5, B6 και B7 (vegetation red edge) είναι κρίσιμα για την εκτίμηση της υγείας της βλάστησης και την ανίχνευση διαφοροποιήσεων στη φυτική κάλυψη, καθώς αποτυπώνουν μεταβάσεις που δεν είναι εμφανείς στο ορατό φάσμα. Τα κανάλια B8 και B8A (NIR και narrow NIR) αξιοποιούνται για τον υπολογισμό δεικτών βλάστησης όπως ο NDVI και για τη διάκριση βλάστησης από άλλες καλύψεις, ενώ τα κανάλια B11 και B12 (SWIR) είναι απαραίτητα για την εκτίμηση της υγρασίας του εδάφους, τη χαρτογράφηση γεωλογικών χαρακτηριστικών και τον εντοπισμό καμένων εκτάσεων. Αντίθετα, απορρίφθηκαν τα κανάλια που δεν προσφέρουν ουσιαστική πληροφορία για την κάλυψη εδάφους: το κανάλι B1 (Coastal aerosol) χρησιμοποιείται κυρίως για ατμοσφαιρικές διορθώσεις και παρακολούθηση παράκτιων περιοχών και δεν συμβάλλει στη διάκριση χρήσεων γης σε ηπειρωτικές περιοχές. Το κανάλι B9 (Water Vapour) προορίζεται για τη μέτρηση της περιεκτικότητας υδρατμών στην ατμόσφαιρα και όχι για τη μελέτη της επιφανειακής κάλυψης. Τέλος, το κανάλι B10 (Cirrus) χρησιμοποιείται για την ανίχνευση αραιών νεφών και δεν περιέχει πληροφορία χρήσιμη για την ανάλυση του εδάφους ή της βλάστησης, ενώ και τα τρία αυτά κανάλια διαθέτουν χαμηλή χωρική ανάλυση (60m), γεγονός που θα υποβάθμιζε τη λεπτομέρεια της τελικής ανάλυσης.



Εικόνα 23: Η αποκοπή των εικόνων

ΚΑΝΑΛΙΑ ΑΡΧΙΚΗΣ ΕΙΚΟΝΑΣ SENTINEL2	ΜΗΚΟΣ ΚΥΜΑΤΟΣ (nm)
2-blue	492.4
3-green	559.8
4-red	664.6
5-vegetation red edge	704.1
6-vegetation red edge	740.5
7-vegetation red edge	782.8
8-nir	832.8
8A-narrow nir	864.7
11-swir	1613.7
12-swir	2202.4

Εικόνα 24: Κανάλια που παρέμειναν

8.5 Εργαλεία και Βιβλιοθήκες

Στην παρούσα υποενότητα καταγράφονται τα εργαλεία και οι βιβλιοθήκες που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία.

8.5.1 Python

Η Python είναι μια υψηλού επιπέδου, γενικού σκοπού και διερμηνευόμενη γλώσσα προγραμματισμού. Ανήκει στην κατηγορία των προστακτικών γλωσσών και υποστηρίζει τόσο το διαδικαστικό όσο και το αντικειμενοστραφές μοντέλο προγραμματισμού. Πρόκειται για μια γλώσσα με δυναμικούς τύπους δεδομένων, η οποία διαθέτει μηχανισμό συλλογής απορριμμάτων (garbage collection) (D. Kuhlman, 2009).

8.5.2 Numpy

Η NumPy¹³ είναι μια βιβλιοθήκη της Python που αναπτύχθηκε για την επεξεργασία και διαχείριση μαθηματικών προβλημάτων μέσω της χρήσης πινάκων. Βρίσκει ευρεία εφαρμογή σε τομείς όπως η γραμμική άλγεβρα, οι μετασχηματισμοί Fourier και οι πράξεις μεταξύ πινάκων. Η βιβλιοθήκη δημιουργήθηκε το 2005 από τον Travis Oliphant και αποτελεί ένα εργαλείο ανοιχτού κώδικα, διαθέσιμο για ελεύθερη χρήση. Ο κύριος λόγος υιοθέτησής της είναι ότι, αν και η Python διαθέτει λίστες που μπορούν να χρησιμοποιηθούν ως πίνακες, αυτές είναι λιγότερο αποδοτικές σε όρους ταχύτητας επεξεργασίας αλλά και ευκολίες εκτέλεσης πράξεων και μετασχηματισμών. Η NumPy, αντίθετα, προσφέρει μια βελτιστοποιημένη δομή πίνακα, εξασφαλίζοντας σημαντικά ταχύτερη εκτέλεση μαθηματικών και αριθμητικών πράξεων σε σύγκριση με τις παραδοσιακές λίστες της Python (Oliphant, 2006).

8.5.3 Scikit-learn

Η scikit-learn¹⁴ (sklearn) είναι μια βιβλιοθήκη ανοιχτού κώδικα για την Python, η οποία αναπτύχθηκε με σκοπό την υλοποίηση αλγορίθμων μηχανικής μάθησης. Παρέχει ένα ευρύ σύνολο εργαλείων για

¹³ <https://numpy.org/>

¹⁴ <https://scikit-learn.org/stable/>

στατιστική μοντελοποίηση και μηχανική μάθηση, όπως ταξινόμηση, παλινδρόμηση, ομαδοποίηση και μείωση διαστάσεων. Η scikit-learn περιλαμβάνει τις ακόλουθες βασικές λειτουργίες: (Hackeling,2017):

- Αλγόριθμους επιβλεπόμενης μάθησης (Supervised learning algorithms)
- Μεθόδους διασταυρούμενης επικύρωσης (Cross-validation)
- Αλγόριθμους μη επιβλεπόμενης μάθησης (Unsupervised learning algorithms)
- Σύνολα δεδομένων μικρού όγκου για δοκιμές (Various toy datasets)
- Τεχνικές εξαγωγής χαρακτηριστικών (Feature extraction)

8.5.4 Geospatial Data Abstraction Library

Η GDAL (Geospatial Data Abstraction Library)¹⁵ είναι μια βιβλιοθήκη ανοιχτού κώδικα που χρησιμοποιείται για την επεξεργασία, ανάγνωση και γραφή γεωχωρικών δεδομένων σε πολλές διαφορετικές μορφές. Αποτελεί ένα βασικό εργαλείο για τη διαχείριση γεωγραφικών δεδομένων, υποστηρίζοντας τόσο ψηφιδωτά/ψηφιογραφικά δεδομένα (π.χ. εικόνες, χάρτες) όσο και δεδομένα διανυσματικών γεωμετριών (π.χ. σημεία, γραμμές, πολυγωνικά σχήματα). Η GDAL παρέχει ένα σύνολο API για την αλληλεπίδραση με γεωχωρικά δεδομένα και προσφέρει υποστήριξη για πλήθος μορφών αρχείων, όπως GeoTIFF, Shapefile, KML, NetCDF, HDF5, και άλλες. Επιπλέον, η GDAL επιτρέπει την εκτέλεση γεωχωρικών αναλύσεων και μετατροπών, καθιστώντας την πολύ χρήσιμη σε εφαρμογές που απαιτούν επεξεργασία γεωγραφικών δεδομένων, όπως GIS (Geographic Information Systems), γεωχωρική ανάλυση και διαχείριση χαρτών. Η βιβλιοθήκη χρησιμοποιείται ευρέως σε εργαλεία και πλατφόρμες GIS και είναι το βασικό εργαλείο πίσω από πολλές δημοφιλείς εφαρμογές και βιβλιοθήκες, όπως το QGIS, το PostGIS και άλλες πλατφόρμες γεωχωρικής ανάλυσης (Ose, 2018).

8.5.5 Semi-Automatic Classification Plugin (SCP)

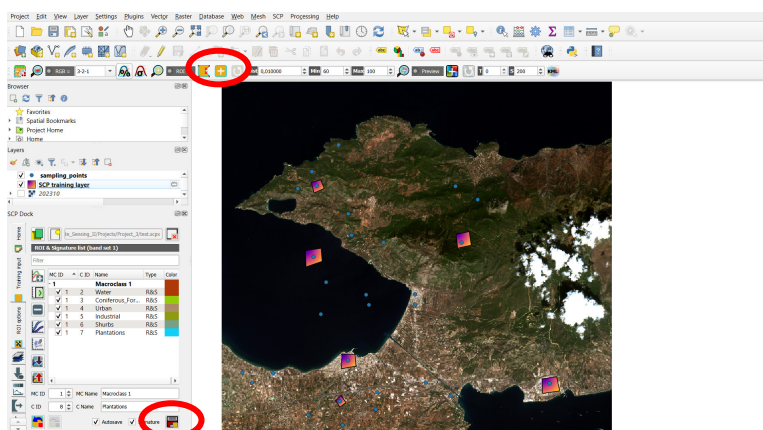
Το Semi-Automatic Classification Plugin (SCP)¹⁶ είναι ένα ισχυρό πρόσθετο του QGIS που διατίθεται δωρεάν και επιτρέπει την ημι-αυτόματη ταξινόμηση δορυφορικών εικόνων, διευκολύνοντας τη χρήση δεδομένων τηλεπισκόπησης για εφαρμογές όπως η χαρτογράφηση χρήσεων γης, η ανάλυση περιβαλλοντικών αλλαγών και η εξαγωγή φασματικών πληροφοριών. Το SCP παρέχει εργαλεία για τη λήψη, προεπεξεργασία και ταξινόμηση δορυφορικών εικόνων από πηγές όπως οι Sentinel-2, Landsat και ASTER, ενώ υποστηρίζει τόσο επιβλεπόμενες όσο και μη επιβλεπόμενες μεθόδους ταξινόμησης μέσω αλγορίθμων μηχανικής μάθησης. Το πρόσθετο αυτό είναι γραμμένο κυρίως σε Python. Στο SCP, αξιοποιούνται βιβλιοθήκες όπως η GDAL (Geospatial Data Abstraction Library) για τη διαχείριση γεωχωρικών δεδομένων, η NumPy για αριθμητικούς υπολογισμούς, η SciPy για επιστημονική ανάλυση, καθώς και η Scikit-learn για την υλοποίηση αλγορίθμων μηχανικής μάθησης. Η ενσωμάτωση του SCP στο QGIS μέσω της PyQGIS API επιτρέπει να αυτοματοποιούνται διαδικασίες, να δημιουργούνται και να ενσωματώνονται scripts σε Python και να αναπτύσσονται προσαρμοσμένες ροές εργασίας για την επεξεργασία τηλεπισκοπικών δεδομένων. Συνεπώς για τις ανάγκες της εργασίας έγινε εγκατάσταση του συγκεκριμένου Plug in.

¹⁵ <https://gdal.org/en/stable/>

¹⁶ <https://plugins.qgis.org/plugins/SemiAutomaticClassificationPlugin/>

8.6 Επιλογή Δεδομένων Εκπαίδευσης

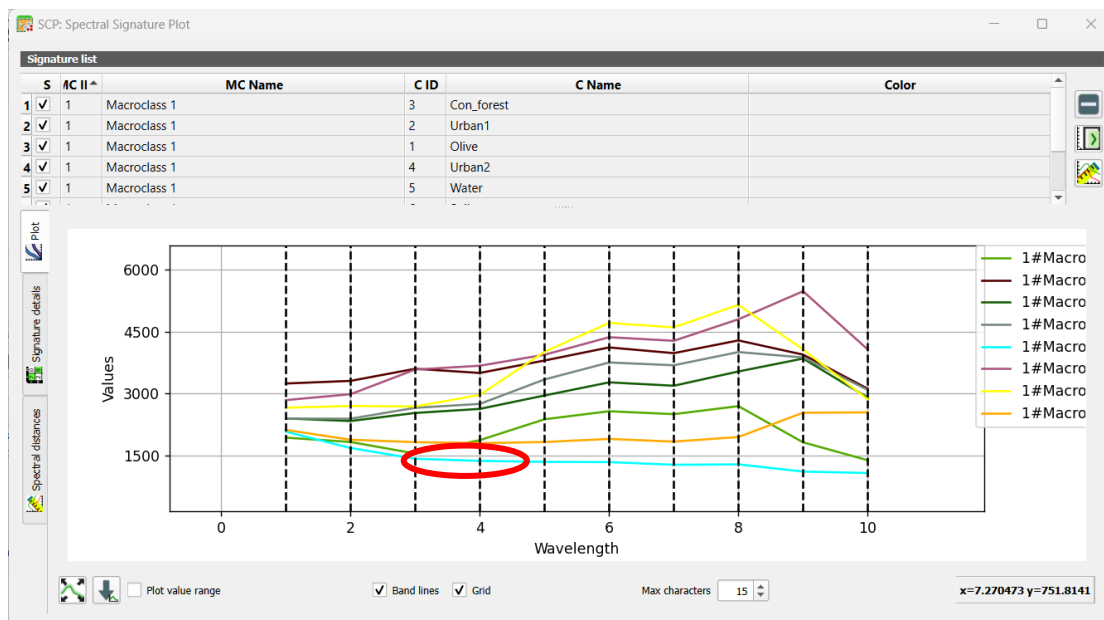
Για την ταξινόμηση των χρήσεων γης χρησιμοποιήθηκε το πρόσθετο Semi-Automatic Classification Plugin (SCP). Αρχικά, πραγματοποιήθηκε η επιλογή των βασικών κατηγοριών χρήσεων γης που εμφανίζονται στην περιοχή μελέτης, όπως γεωργικές εκτάσεις, δάση, αστικές περιοχές, υδάτινες επιφάνειες κλπ. Για κάθε κατηγορία δημιουργήθηκαν δεδομένα εκπαίδευσης με τη χάραξη πολυγώνων απευθείας πάνω στην εικόνα Sentinel-2. Συγκεκριμένα, για κάθε τύπο χρήσης γης σχεδιάστηκαν δύο πολύγωνα, περίπου ίδιου μεγέθους, ώστε να υπάρχει καλύτερη αντιπροσώπευση της ποικιλομορφίας μέσα σε κάθε κατηγορία. (Με τα εργαλεία που είναι στους κόκκινους κύκλους) Η δημιουργία των πολυγώνων έγινε με στόχο να επιλέγονται όσο το δυνατόν πιο ομοιογενείς περιοχές ως προς το χρώμα και την υφή, ώστε να αυξηθεί η ακρίβεια της ταξινόμησης. Τα πολύγωνα αυτά ορίστηκαν ως δείγματα εκπαίδευσης (training samples) μέσα από τη λειτουργία «Training Input» του SCP. Κατά τη διαδικασία δημιουργίας των πολυγώνων δόθηκε προσοχή ώστε να καλύπτονται διαφορετικά σημεία της εικόνας και να αποφεύγονται περιοχές με σκιές ή σύννεφα που θα μπορούσαν να προκαλέσουν σφάλματα στην ταξινόμηση. Αφού ολοκληρώθηκε η συλλογή των δεδομένων εκπαίδευσης, τα πολύγωνα αυτά χρησιμοποιήθηκαν για την εκπαίδευση του αλγορίθμου ταξινόμησης, με σκοπό τη δημιουργία ενός χάρτη χρήσεων γης της περιοχής μελέτης.



Εικόνα 25: Δημιουργία Πολυγώνων

0	- Unclassified
3	- Con_forest
2	- Urban1
1	- Olive
4	- Urban2
5	- Water
6	- Soil
7	- Vegetations
8	- Burnt

Εικόνα 26: Πολύγωνα που δημιουργήθηκαν



Εικόνα 27: Στατιστικά στοιχεία δεδομένων εκπαίδευσης

Η παραπάνω εικόνα προέρχεται από το εργαλείο Spectral Signature Plot του Semi-Automatic Classification Plugin (SCP) στο QGIS και παρουσιάζει τις φασματικές υπογραφές των διαφορετικών κατηγοριών κάλυψης γης που επιλέχθηκαν στην περιοχή μελέτης. Συγκεκριμένα, για κάθε κατηγορία εδάφους εμφανίζεται η μέση τιμή της ανακλώμενης ακτινοβολίας σε διαφορετικά μήκη κύματος, όπως ανιχνεύονται από τα κανάλια του δορυφόρου Sentinel-2. Η οριζόντια άξονας (Wavelength) αντιπροσωπεύει τα διαφορετικά φασματικά κανάλια, ενώ ο κάθετος άξονας (Values) απεικονίζει τις τιμές ανάκλασης (ανακλώμενη ενέργεια).

Στο διάγραμμα απεικονίζονται οι φασματικές καμπύλες για οκτώ συνολικά κατηγορίες χρήσης γης, κάθε μία με διαφορετικό χρώμα γραμμής. Οι κατηγορίες και τα αντίστοιχα χρώματα είναι τα εξής:

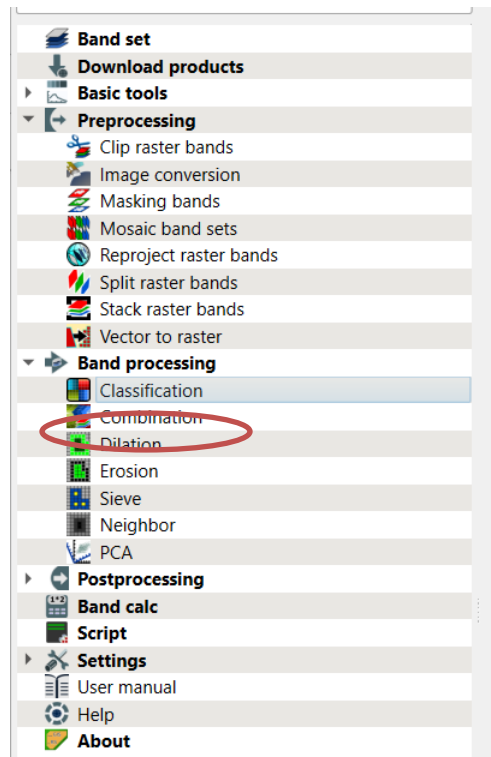
- 0 - Unclassified: Μαύρο
- 3 - Coniferous forest (Δάσος Κωνοφόρων): Πράσινο
- 2 - Urban1 (Αστική περιοχή 1): Καφέ
- 1 - Olive (Ελαιώνας): Σκούρο Πράσινο
- 4 - Urban2 (Αστική περιοχή 2): Γκρι
- 5 - Water (Υδάτινες επιφάνειες): Γαλάζιο
- 6 - Soil (Γυμνό έδαφος): Μωβ
- 7 - Vegetations (Γενική Βλάστηση): Κίτρινο
- 8 - Burnt (Καμένη έκταση): Πορτοκαλί

Η απεικόνιση αυτή είναι ιδιαίτερα χρήσιμη γιατί αναδεικνύει τις διαφορές ανάμεσα στις κατηγορίες ως προς την ανακλαστικότητα σε διάφορες φασματικές περιοχές. Για παράδειγμα, η κατηγορία "Water" εμφανίζει πολύ χαμηλές τιμές ανάκλασης σε όλα τα μήκη κύματος, ενώ η "Vegetation" και τα δάση ("Coniferous forest") παρουσιάζουν χαρακτηριστική υψηλή ανακλαστικότητα στο εγγύς υπέρυθρο (NIR) κανάλι. Αυτές οι διαφορές αξιοποιούνται στη διαδικασία της ταξινόμησης, ώστε να γίνει ο διαχωρισμός των τύπων χρήσης γης με βάση τις φασματικές ιδιότητες κάθε κατηγορίας.

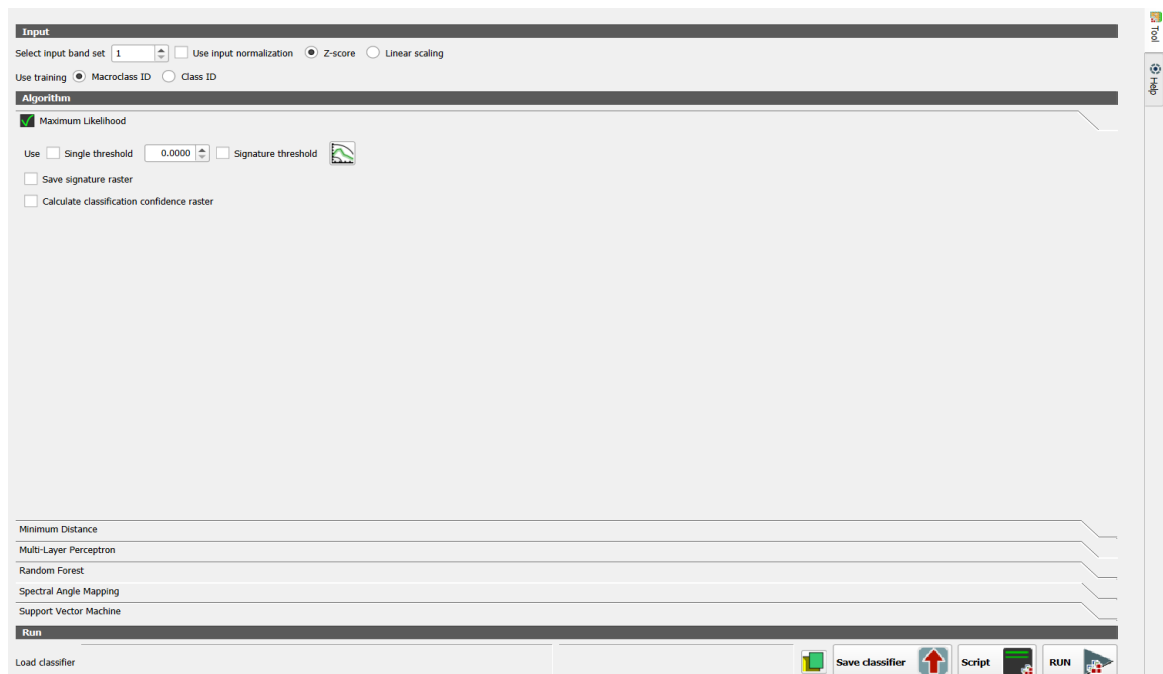
Τα διαγράμματα των φασματικών υπογραφών που προκύπτουν από το Semi-Automatic Classification Plugin (SCP) είναι ιδιαίτερα σημαντικά για την ταξινόμηση δορυφορικών εικόνων, καθώς επιτρέπουν την ταυτοποίηση και διάκριση διαφορετικών τύπων καλύψεων γης (όπως βλάστηση, νερό, έδαφος και αστικές περιοχές) με βάση τα φασματικά τους χαρακτηριστικά. Το SCP λειτουργεί με βάση τις φασματικές υπογραφές, δηλαδή τις μοναδικές ανακλάσεις των αντικειμένων σε διάφορα μήκη κύματος, οι οποίες καταγράφονται από τους αισθητήρες των δορυφόρων, όπως οι Sentinel-2 και Landsat. Μέσω των διαγραμμάτων αυτών, οι χρήστες μπορούν να αναλύσουν και να συγκρίνουν τις φασματικές αποκρίσεις διαφορετικών υλικών, καθιστώντας δυνατή τη επιβλεπόμενη κατηγοριοποίηση των εικόνων. Η διαδικασία αυτή επιτρέπει την ακριβή εκπαίδευση των μοντέλων ταξινόμησης, καθώς οι φασματικές υπογραφές χρησιμοποιούνται ως δεδομένα αναφοράς για την αναγνώριση παρόμοιων περιοχών στην εικόνα. Επιπλέον, τα διαγράμματα των φασματικών υπογραφών βοηθούν στην αξιολόγηση της ποιότητας και της συνέπειας των δεδομένων, επιτρέποντας τη βελτίωση των αλγορίθμων και την αύξηση της ακρίβειας των ταξινομήσεων. Συνεπώς, η χρήση των φασματικών υπογραφών στο SCP αποτελεί τη βάση για μια αξιόπιστη και συστηματική ανάλυση τηλεπισκοπικών δεδομένων, διευκολύνοντας τη χαρτογράφηση της γης και την παρακολούθηση περιβαλλοντικών αλλαγών.

8.7 Κατηγοριοποιήσεις

Στην παρούσα υποενότητα καταγράφονται τα αποτελέσματα κατηγοριοποιήσεων που πραγματοποιήθηκαν με χρήση του SCP, και αξιολογείται η ποιότητα των αποτελεσμάτων αυτών. Για την μέγιστη πιθανοφάνεια αφού δημιουργήθηκαν τα δεδομένα εκπαίδευσης (training polygons) για κάθε κατηγορία χρήσης γης, επιλέχθηκε από το μενού του SCP η ενότητα Classification. Στο παράθυρο Classification Algorithm ορίστηκε ως μέθοδος ταξινόμησης η Maximum Likelihood. Πριν την εκτέλεση, ελέγχθηκαν οι παράμετροι ώστε να χρησιμοποιηθεί η πλήρης φασματική πληροφορία από τα επιλεγμένα κανάλια Sentinel-2. Έπειτα, πραγματοποιήθηκε η εκτέλεση της ταξινόμησης, με αποτέλεσμα να παραχθεί χάρτης όπου κάθε pixel ταξινομήθηκε στην κατηγορία για την οποία εμφάνιζε τη μεγαλύτερη στατιστική πιθανότητα ανήκειν, με βάση τα δείγματα εκπαίδευσης. Στην συνέχεια για τη Φασματική γωνία στο ίδιο παράθυρο Classification Algorithm επιλέχθηκε η μέθοδος Spectral Angle Mapping. Ο συγκεκριμένος αλγόριθμος απαιτεί τον υπολογισμό του φασματικού διανύσματος για κάθε κατηγορία εκπαίδευσης. Με την εκτέλεση της ταξινόμησης, δημιουργήθηκε θεματικός χάρτης, όπου κάθε pixel ταξινομήθηκε στην κατηγορία με την ελάχιστη φασματική γωνία σε σχέση με τις υπογραφές των δεδομένων εκπαίδευσης. Για την ταξινόμηση με ελάχιστη απόσταση, και πάλι μέσω του παραθύρου Classification Algorithm του SCP, επιλέχθηκε ο αλγόριθμος Minimum Distance. Η μέθοδος αυτή απαιτεί τον υπολογισμό της ευκλείδειας απόστασης κάθε pixel από τα μέσες φασματικές τιμές των κατηγοριών εκπαίδευσης. Δεν απαιτήθηκαν ειδικές ρυθμίσεις πέρα από την επιβεβαίωση ότι χρησιμοποιούνται τα κατάλληλα φασματικά κανάλια. Με την εκτέλεση της διαδικασίας, προέκυψε ταξινομημένος χάρτης όπου κάθε pixel ανατέθηκε στην κοντινότερη, βάσει απόστασης, κατηγορία. Πριν από κάθε ταξινόμηση, τα δεδομένα εκπαίδευσης επανασυνδέθηκαν (linked) με τη φασματική εικόνα. Σε όλες τις ταξινομήσεις επιλέχθηκε η επιλογή Create Classification Preview για έλεγχο των αποτελεσμάτων σε μικρότερη ανάλυση, πριν την πλήρη εκτέλεση. Μετά από κάθε ταξινόμηση δημιουργήθηκε αρχείο raster, το οποίο αποθηκεύτηκε για μεταγενέστερη αξιολόγηση



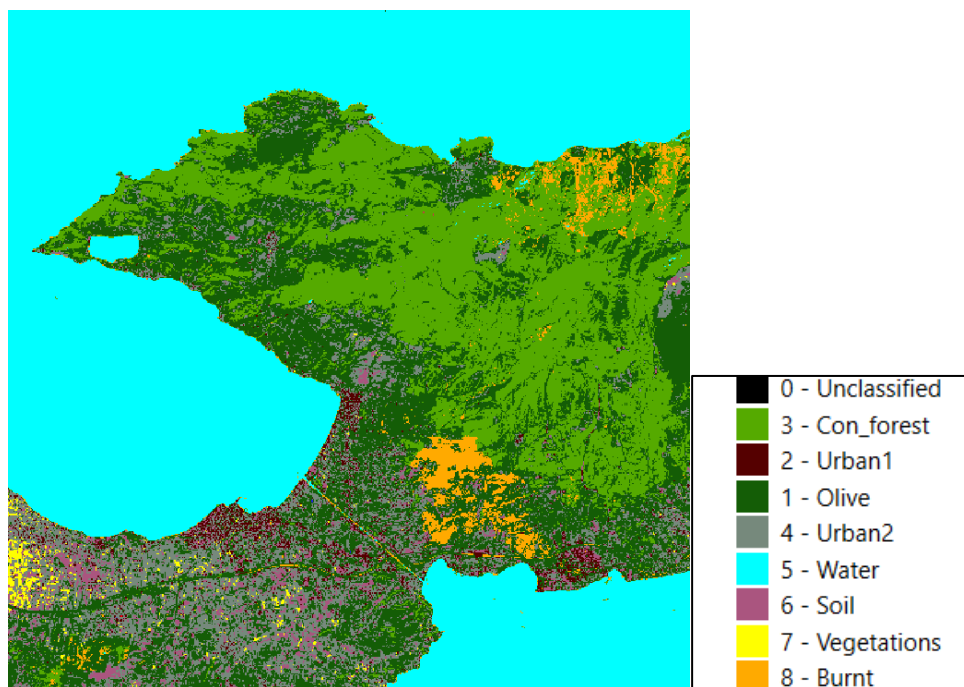
Εικόνα 28: Επιλογή Classification στο SCP



Εικόνα 29: Επιλογή Αλγορίθμου

8.7.1 Κατηγοριοποίηση ελάχιστης απόστασης

Στην Εικόνα 30 απεικονίζεται οπτικά η κατηγοριοποίηση που επιτεύχθηκε με τη μέθοδο της ελάχιστης απόστασης.



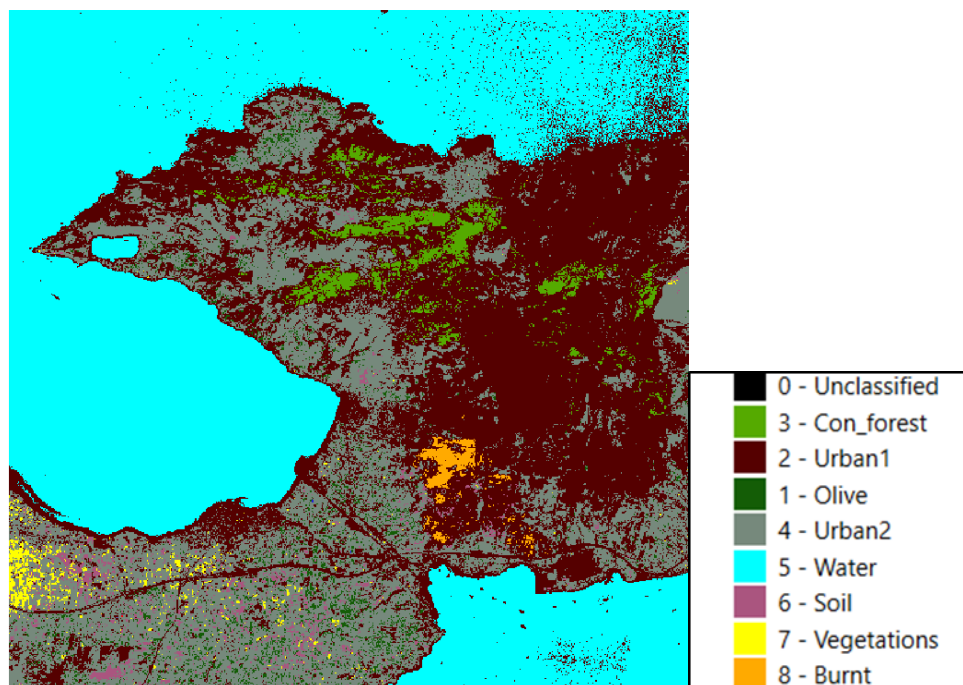
Εικόνα 30: Ταξινόμηση ελάχιστης απόστασης

Αρχικά φαίνεται ότι το αποτέλεσμα ανταποκρίνεται σε ένα βαθμό στην πραγματικότητα με βάση τα πολύγωνα εκπαίδευσης που ορίσαμε. Το μεγαλύτερο ορεινό κομμάτι των Γερανείων φαίνεται να είναι κωνοφόρο δάσος (Con_forest) και στους πρόποδες να υπάρχουν ελαιώνες (Olive). Επίσης ελαιώνες και γυμνό έδαφος υπάρχουν περίξ του αστικού ιστού, τόσο του συνεχούς όσο και του ασυνεχούς¹⁷. Χαρακτηριστικές είναι και οι περιοχές που υπάρχουν καλλιέργειες αλλά και οι καμένες περιοχές. Ωστόσο αν αναλύσουμε πιο ενδελεχώς την εικόνα θα εντοπιστούν ορισμένα σημαντικά σφάλματα. Πιο συγκεκριμένα υπάρχει σύγχυση του συνεχούς και του ασυνεχούς ιστού καθώς οι φασματικές τους υπογραφές είναι πολύ κοντά. Επίσης κάποια γυμνά εδάφη φαίνονται ως ελαιώνες καθώς η αραιή συγκέντρωση των ελαιώνων κάνει την φασματική υπογραφή να βρίσκεται κοντά στο γυμνό έδαφος. Ένα ακόμα σημαντικό σφάλμα φαίνεται στον αστικό ιστό, όπου φαίνονται καλλιέργειες ενώ δεν υπάρχουν. Η αιτία του συνόλου των σφαλμάτων θα μπορούσε να αναχθεί στο γεγονός ότι υπάρχει ανομοιογένεια ή ανομοιομορφία των πολυγώνων εκπαίδευσης, καθώς και απουσίας εξειδικευμένης κατηγοριοποίησης για κάθε διαφορετική έκφανση των κατηγοριών.

¹⁷ Ο συνεχής αστικός ιστός περιλαμβάνει περιοχές όπου η επιφάνεια καλύπτεται σε ποσοστό μεγαλύτερο του 80% από τεχνητές κατασκευές (κτήρια, δρόμοι, κ.λπ.). Ο ασυνεχής αστικός ιστός περιλαμβάνει περιστατικές περιοχές γύρω από αστικά κέντρα καθώς και κάποιους οικισμούς σε αγροτικές περιοχές με κάλυψη επιφάνειας από τεχνητές κατασκευές μικρότερη του 80%.

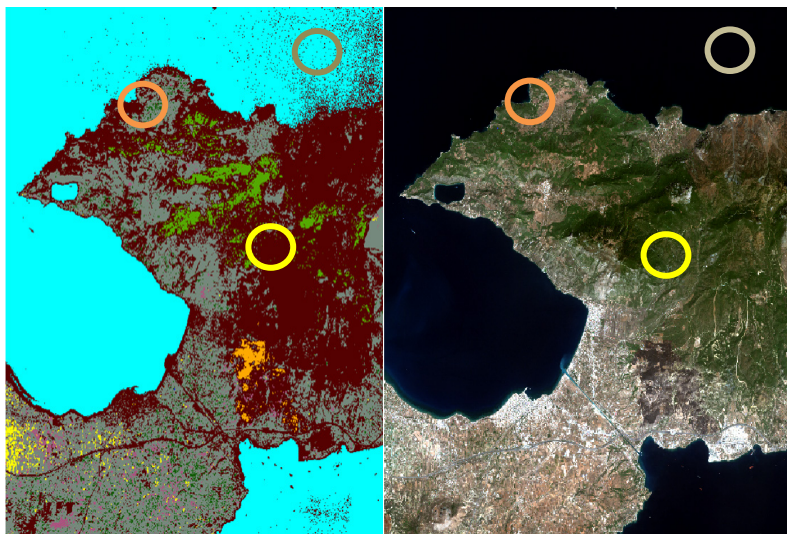
8.7.2 Κατηγοριοποίηση Μέγιστης Πιθανοφάνειας

Στην Εικόνα 31 απεικονίζεται οπτικά η κατηγοροποίηση που επιτεύχθηκε με τη μέθοδο της Μέγιστης Πιθανοφάνειας.



Εικόνα 31: Ταξινόμηση μέγιστης πιθανοφάνειας

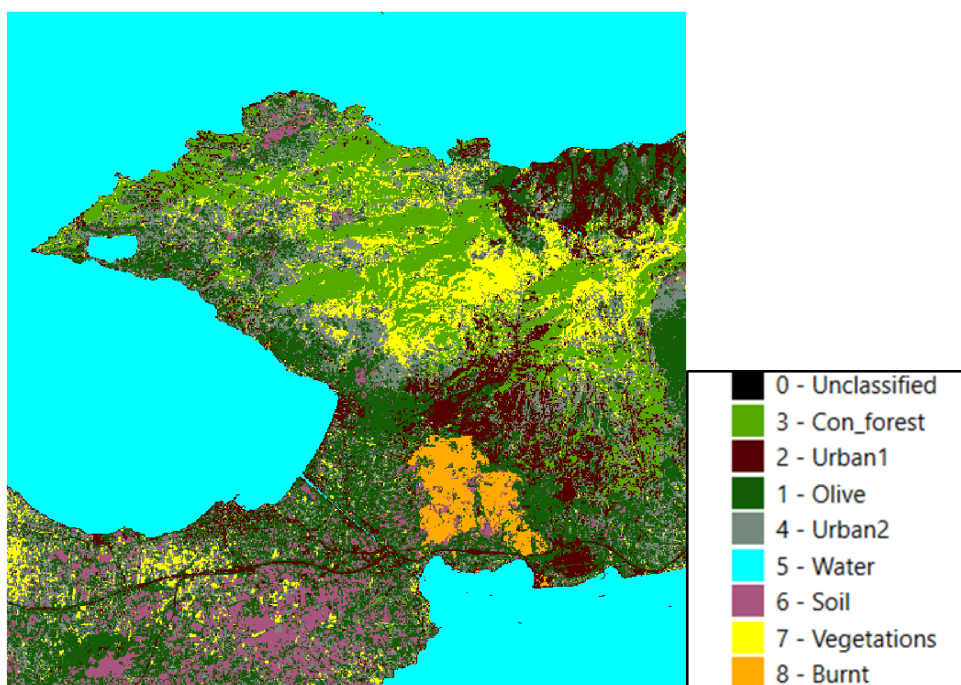
Στην παραπάνω κατηγοριοποίηση παρατηρείται ότι, το αποτέλεσμα γενικά αποκλίνει σε μεγάλο βαθμό από την αρχική εικόνα. Πιο συγκεκριμένα, μεγάλες εκτάσεις που θα έπρεπε να ταξινομηθούν στην κατηγορία γυμνό έδαφος, ελαιώνες και κάποια δάση ταξινομήθηκαν στην κατηγορία Αστικός Ιστός. Αυτό συμβαίνει διότι, η μέθοδος της Μέγιστης Πιθανοφάνειας ταξινομεί με βάση την στατιστική πιθανότητα να ανήκει ένα εικονοστοιχείο σε μία κατηγορία, λαμβάνοντας υπόψη το μέσο και τη διασπορά της κατηγορίας. Έτσι, ένα στοιχείο είναι πιο πιθανό να ενταχθεί σε μια κατηγορία που στα δεδομένα εκπαίδευσης εμφανίζει μεγάλη διασπορά, παρά σε άλλες κατηγορίες που έχουν με μικρότερη διασπορά. Όπως φαίνεται και στα ιστογράμματα, η κατηγορία Αστικός Ιστός έχει μεγάλη διασπορά ενώ π.χ. το γυμνό έδαφος (κίτρινο στο ιστόγραμμα) παρουσιάζει μικρότερη. Καθώς λοιπόν οι δύο παραπάνω κατηγορίες είναι «φασματικά γειτονικές» αλλά η μια έχει μεγαλύτερη «φασματική βαρύτητα» από την άλλη στην συγκεκριμένη μέθοδο, όλα τα εικονοστοιχεία που ανήκουν στην κατηγορία γυμνό έδαφος ταξινομούνται ως αστικός ιστός. Σίγουρα μερίδιο ευθύνης υπάρχει και στο γεγονός πως υπάρχει αστοχία στην κατασκευή πολυγώνων. Σημαντικά σφάλματα εμφανίζονται και στη θάλασσα, όπου κάποια στοιχεία της απεικονίζονται ως αστικός ιστός, το γυμνό έδαφος που απεικονίζεται ως αστικός ιστός αλλά και το κωνοφόρο δάσος που απεικονίζεται ως αστικός ιστός. Στην περίπτωση αυτή, ως αστικός ιστός νοείται και ο συνεχής και ο μη καθώς παρουσιάζουν παρόμοια συμπεριφορά στα στατιστικά στοιχεία και στη φασματική υπογραφή.



Εικόνα 32: Σημαντικά σφάλματα (με διαφορετικό χρώμα κάθε περίπτωση)

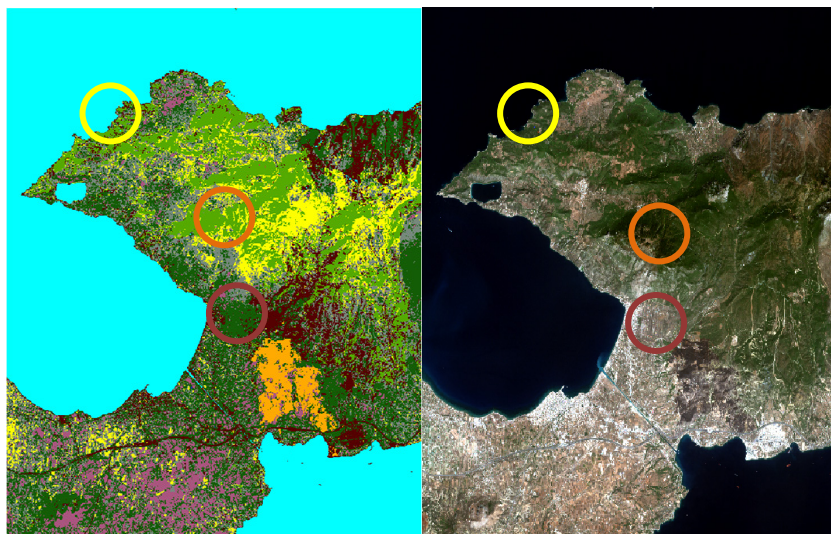
8.7.3 Κατηγοριοποίηση φασματικής γωνίας

Στην Εικόνα 32 απεικονίζεται οπτικά η κατηγοριοποίηση που επιτεύχθηκε με τη μέθοδο της φασματικής γωνίας.



Εικόνα 33. Κατηγοριοποίηση φασματικής γωνίας

Στην κατηγοριοποίηση αυτή, παρατηρείται ότι, το αποτέλεσμα γενικά είναι κοντά στην αρχική εικόνα όπως και στην μέθοδο της ελάχιστης απόστασης. Μεγάλο σφάλμα υπάρχει σε κάποια δασώδη περιοχή η οποία έχει κατηγοριοποιηθεί ως καλλιέργεια. Κάποια γυμνά εδάφη επίσης τα παρουσιάζει ως αστικό ιστό. Παρ' όλα τα σφάλματα, η μέθοδος αυτή είναι αρκετά αποτελεσματική ως προς την λεπτομέρεια καθώς, ταξινομήσε ορθά το κύριο οδικό δίκτυο και διατήρησε τα σχήματα.



Εικόνα 34: Σημαντικά σφάλματα (με διαφορετικό χρώμα κάθε περίπτωση)

Στις δύο τελευταίες περιπτώσεις δεν χρειάστηκε να αυξηθεί η μεγέθυνση (ζουμ) για να εντοπιστούν κάποια από τα σημαντικά σφάλματα.

Όσον αφορά την ταξινόμηση που έδωσε τα καλύτερα αποτελέσματα πρέπει να αναφερθεί πως, η συγκεκριμένη κατάσταση ισχύει μονάχα για τα συγκεκριμένα πολύγωνα εκπαίδευσης. Διαφορετικά πολύγωνα, θα είχαν ίσως διαφορετικά αποτελέσματα. Στην προκειμένη περίπτωση λοιπόν τα καλύτερα αποτελέσματα τα δίνει η μέθοδος της φασματικής γωνίας καθώς, οι άλλες δύο μέθοδοι έχουν την τάση να γενικεύουν «επιπόλαια» την διαδικασία της ταξινόμησης λόγω της φύσης των αλγορίθμων τους. Εδώ παρατηρείται ότι η μέθοδος φασματικής γωνίας κατάφερε να αποτυπώσει ακόμα και το οδικό δίκτυο αλλά και διατηρήθηκαν τα σχήματα. Η μέθοδος της φασματικής γωνίας είναι αποδοτικότερη όταν τα δεδομένα είναι πιο αυστηρά καθώς, σε πολλές διαστάσεις, αυξάνεται η πολυπλοκότητα της κάθε γωνίας. Πιο αυστηρά σημαίνει ότι, απαιτείται τα πολύγωνα εκπαίδευσης να είναι πλήρως αντιπροσωπευτικά και ομοιογενή, να περιλαμβάνουν μόνο περιοχές με καθαρή, σαφώς καθορισμένη χρήση γης και να αποφεύγονται περιοχές με μικτές καλύψεις, σκιές ή άλλες ασαφείς συνθήκες. Η αυστηρότητα από την πλευρά του χρήστη διασφαλίζει ότι οι φασματικές υπογραφές που δημιουργούνται για κάθε κατηγορία είναι καθαρές και ευδιάκριτες, γεγονός που είναι κρίσιμο για τη SAM, καθώς η μέθοδος βασίζεται στη μέτρηση της γωνίας μεταξύ των φασματικών διανυσμάτων και όχι στην απόλυτη τιμή της ανακλώμενης ακτινοβολίας. Σε περιπτώσεις όπου τα δεδομένα εκπαίδευσης περιέχουν θόρυβο ή αναμειγνύονται διαφορετικών τύπων κάλυψης, το σχήμα της φασματικής υπογραφής αλλοιώνεται, γεγονός που οδηγεί σε χαμηλότερη ακρίβεια ταξινόμησης. Επομένως, η επιτυχία της μεθόδου εξαρτάται σε μεγάλο βαθμό από την προσεκτική και αυστηρή επιλογή των δειγμάτων εκπαίδευσης από τον χρήστη. Επίσης, η συγκεκριμένη μέθοδος αφήνει βαθμούς ελευθερίας στην ένταση. Για παράδειγμα στις δύο διαστάσεις, η μέθοδος είναι επιρρεπής στο να συγχέει κατηγορίες οι οποίες είναι η μια πίσω από την άλλη καθώς, και οι δύο σχηματίζουν την ίδια φασματική γωνία με το τυχαίο προς έλεγχο εικονοστοιχείο. Όταν αναφέρεται ότι η μέθοδος της φασματικής γωνίας αφήνει βαθμούς ελευθερίας στην ένταση, εννοείται ότι κατά την ταξινόμηση δεν λαμβάνεται υπόψη το απόλυτο μέγεθος της ανακλώμενης ακτινοβολίας, αλλά μόνο η διεύθυνση του φασματικού διανύσματος στον πολυδιάστατο φασματικό χώρο. Ο

αλγόριθμος βασίζεται στη μέτρηση της γωνίας, γνωστής ως Spectral Angle, που σχηματίζει το διάνυσμα ενός pixel με το διάνυσμα της κάθε κλάσης εκπαίδευσης. Αυτό σημαίνει ότι δύο κατηγορίες που έχουν παρόμοιο σχήμα φασματικής υπογραφής αλλά διαφορετικά επίπεδα φωτεινότητας ή έντασης μπορούν να σχηματίσουν την ίδια φασματική γωνία σε σχέση με ένα τυχαίο pixel, με αποτέλεσμα να υπάρχει πιθανότητα σύγχυσης στην ταξινόμηση. Δεν αναφερόμαστε σε κατακόρυφες ή άλλες γεωμετρικές γωνίες, αλλά σε τη φασματική γωνία που προκύπτει από την κατεύθυνση των διανυσμάτων στο φασματικό χώρο. Το φαινόμενο αυτό είναι γνωστό ως spectral confusion και αποτελεί χαρακτηριστικό περιορισμό της μεθόδου SAM, ιδίως όταν τα δεδομένα εκπαίδευσης δεν έχουν σαφώς διακριτές φασματικές υπογραφές.

8.8 Αξιολόγηση – Πολύγωνα Ελέγχου

Στη συνέχεια δημιουργήθηκαν πολύγωνα ελέγχου έτσι ώστε να μπορέσει να γίνει η αξιολόγηση των ταξινομήσεων που προηγήθηκαν. Τα πολύγωνα αυτά ήταν διαφορετικά από τα πολύγωνα εκπαίδευσης που χρησιμοποιήθηκαν. Στις επόμενες παραγράφους παρουσιάζονται τα αποτελέσματα από την αξιολόγηση με βάση τα πολύγωνα ελέγχου.

8.8.1 Πίνακας Σύγχυσης Ελάχιστης Απόστασης

Ο Πίνακας 3 παρουσιάζει τον πίνακα σύγχυσης για την κατηγοριοποίηση με βάση την ελάχιστη απόσταση.

ERROR MATRIX [pixel count]									
	> Reference								
V_Classified	1	2	3	4	5	6	7	8	Total
1	90	10	0	0	0	0	0	0	100
2	0	55	0	22	0	0	0	0	77
3	0	0	97	0	0	0	0	0	97
4	1	24	0	40	0	24	0	0	89
5	0	0	0	0	114	0	0	0	114
6	0	1	0	19	0	77	0	0	97
7	0	0	0	1	0	0	93	0	94
8	0	0	0	0	0	0	0	78	78
Total	91	90	97	82	114	101	93	78	746
Overall accuracy [%] = 86.3271									

Πίνακας 3: Πίνακας σύγχυσης ελάχιστης απόστασης

Αφού εξάχθηκε ο πίνακας σύγχυσης είναι σημαντικό να παρουσιασθούν οι ακρίβειες του χρήστη και του παραγωγού για να εξετάσουμε ανά κατηγορία τα αποτελέσματα. Ως *ακρίβεια παραγωγού* (producer's accuracy) για μία κατηγορία K ορίζεται η ακρίβεια της κατηγοριοποίησης από την οπτική του σχεδιαστή/παρόχου του χάρτη και δίνεται από τον μαθηματικό τύπο:

$$\text{ΑκρίβειαΠαραγωγού}(K) = \frac{\text{Πλήθος των περιοχών κατηγορίας } K \text{ που κατατάχθηκαν σωστά}}{\text{Συνολικό πλήθος περιοχών της κατηγορίας } K \text{ στον χάρτη}}$$

Από την άλλη πλευρά, ως *ακρίβεια χρήστη* (user's accuracy) για μια κατηγορία K ορίζεται η ακρίβεια της κατηγοριοποίησης από την οπτική του χρήστη του χάρτη και δίνεται από τον μαθηματικό τύπο:

$$\text{ΑκρίβειαΧρήστη}(K) = \frac{\text{Πλήθος των περιοχών κατηγορίας } K \text{ που κατατάχθηκαν σωστά}}{\text{Συνολικό πλήθος περιοχών που κατατάχθηκαν στην κατηγορία } K}$$

Ακρίβειες ταξινόμησης παραγωγού:

- 1 (Ελαιώνας): $90/91 \cdot 100\% = 98,90\%$
- 2 (Συνεχής αστικός ιστός): $55/90 \cdot 100\% = 61,11\%$
- 3 (Νερό): $97/97 \cdot 100\% = 100\%$
- 4 (Κωνοφόρο δάσος): $40/82 \cdot 100\% = 48,78\%$
- 5 (Ασυνεχής αστικός ιστός): $114/114 \cdot 100\% = 100\%$
- 6 (Καλλιέργειες): $77/101 \cdot 100\% = 79\%$
- 7 (Γυμνό έδαφος): $93/93 \cdot 100\% = 100\%$
- 8 (Καμένη περιοχή): $78/78 \cdot 100\% = 100\%$

Στην κατηγορία 1 (ελαιώνας) παρατηρείται ότι ένα εικονοστοιχείο μόνο δεν ταξινομήθηκε σωστά και εντάχθηκε εσφαλμένα στην κατηγορία 4 (κωνοφόρο δάσος). Στην κατηγορία 2 (συνεχής αστικός ιστός) υπήρξαν 34 εικονοστοιχεία που ταξινομήθηκαν σε διαφορετική κατηγορία όπως και στην κατηγορία 4 (κωνοφόρο δάσος) υπήρχε μεγάλος αριθμός εικονοστοιχείων που ταξινομήθηκε σε διαφορετική κατηγορία. Τέλος στην κατηγορία 6 (καλλιέργειες), από τα 101 εικονοστοιχεία που είχαν οριστεί ως καλλιέργειες, τα 24 δεν ταξινομήθηκαν σωστά.

Ακρίβειες ταξινόμησης χρήστη:

- 1 (Ελαιώνας): $90/100 \cdot 100\% = 90\%$
- 2 (Συνεχής αστικός ιστός): $55/77 \cdot 100\% = 71,14\%$
- 3 (Νερό): $97/97 \cdot 100\% = 100\%$
- 4 (Κωνοφόρο δάσος): $40/89 \cdot 100\% = 44,94\%$
- 5 (Ασυνεχής αστικός ιστός): $114/114 \cdot 100\% = 100\%$
- 6 (Καλλιέργειες): $77/97 \cdot 100\% = 79,38\%$
- 7 (Γυμνό έδαφος): $93/94 \cdot 100\% = 98,93\%$
- 8 (Καμένη περιοχή): $78/78 \cdot 100\% = 100\%$

Στην κατηγορία 1 παρατηρείται ότι 10 εικονοστοιχεία δεν ταξινομήθηκαν σωστά. Στην κατηγορία 2 υπήρξαν 22 εικονοστοιχεία που κατατάχθηκαν σε διαφορετική κατηγορία, ενώ στην κατηγορία 4 (κωνοφόρο δάσος) 44 εικονοστοιχεία ταξινομήθηκαν σε διαφορετική κατηγορία. Στην κατηγορία 6, 77 από τα 97 εικονοστοιχεία εντάχθηκαν στην σωστή κατηγορία και, τέλος, στην κατηγορία 7 από τα 93 εικονοστοιχεία μόνο ένα δεν ταξινομήθηκε σωστά.

Η συνολική ακρίβεια δείχνει το ποσοστό των εικονοστοιχείων που ταξινομήθηκε σωστά από το σύνολο των εικονοστοιχείων των πολύγωνων εκπαίδευσης και το ποσοστό στον συγκεκριμένο τρόπο κατηγοριοποίησης είναι 86.32%, δηλαδή πολύ ικανοποιητικό ποσοστό.

8.8.2 Πίνακας Σύγχυσης Μέγιστης Πιθανοφάνειας

Ο Πίνακας 4 παρουσιάζει τον πίνακα σύγχυσης για την κατηγοριοποίηση με βάση τη μέγιστη πιθανοφάνεια.

ERROR MATRIX [pixel count]									
	> Reference								
V_Classified	1	2	3	4	5	6	7	8	Total
1	55	0	0	0	0	0	0	0	55
2	0	83	0	9	0	0	1	0	93
3	0	0	97	0	0	0	0	0	97
4	36	7	0	72	0	37	0	0	152
5	0	0	0	0	114	0	0	0	114
6	0	0	0	1	0	64	0	0	65
7	0	0	0	0	0	0	92	0	92
8	0	0	0	0	0	0	0	78	78
Total	91	90	97	82	114	101	93	78	746
Overall accuracy [%] = 87.8016									

Πίνακας 4: Πίνακας σύγχυσης μέγιστης πιθανοφάνειας

Στη συνέχεια αναλύονται τα αποτελέσματα, με βάση την ακρίβεια κατηγοριοποίησης παραγωγού, χρήση αλλά και της συνολικής ακρίβειας.

Ακρίβειες κατηγοριοποίησης παραγωγού:

- 1 (Ελαιώνας): $55/91 \cdot 100\% = 60.43\%$
- 2 (Συνεχής αστικός ιστός): $83/90 \cdot 100\% = 92.22\%$
- 3 (Νερό): $97/97 \cdot 100\% = 100\%$
- 4 (Κωνοφόρο Δάσος): $72/82 \cdot 100\% = 87.80\%$
- 5 (Ασυνεχής αστικός ιστός): $114/114 \cdot 100\% = 100\%$
- 6 (Καλλιέργειες): $64/101 \cdot 100\% = 63.36\%$
- 7 (Γυμνό έδαφος): $92/93 \cdot 100\% = 98.92\%$
- 8 (Καμένη περιοχή): $78/78 \cdot 100\% = 100\%$

Στην ανωτέρω καταγραφή φαίνεται αναλυτικά πόσα εικονοστοιχεία από κάθε κατηγορία έχουν ταξινομηθεί σωστά. Τρεις κατηγορίες παρουσιάζουν απόλυτη επιτυχία στην κατηγοριοποίηση, μετά οι υπόλοιπες κυμαίνονται σε υψηλά ποσοστά, άνω του 60%.

Ακρίβειες κατηγοριοποίησης χρήστη:

- 1 (Ελαιώνας): $55/55 \cdot 100\% = 100\%$
- 2 (Συνεχής αστικός ιστός): $83/93 \cdot 100\% = 89.94\%$
- 3 (Νερό): $97/97 \cdot 100\% = 100\%$
- 4 (Κωνοφόρο Δάσος): $72/152 \cdot 100\% = 47,36\%$
- 5 (Ασυνεχής αστικός ιστός): $114/114 \cdot 100\% = 100\%$
- 6 (Καλλιέργειες): $64/65 \cdot 100\% = 98.46\%$
- 7 (Γυμνό έδαφος): $92/92 \cdot 100\% = 100\%$
- 8 (Καμένη περιοχή): $78/78 \cdot 100\% = 100\%$

Από πλευράς χρήστη, οι περισσότερες κατηγορίες είναι σχεδόν τέλεια ταξινομημένες. Η κατηγορία του κωνοφόρου δάσους φαίνεται να έχει μικρό ποσοστό όπου από τα 152, μόνον τα 72 ταξινομήθηκαν σωστά.

Η συνολική ακρίβεια στην παρούσα ταξινόμηση είναι 87.80%, δηλαδή παραμένει σε υψηλό επίπεδο.

8.8.3 Πίνακας Σύγχυσης Φασματικής Γωνίας

Ο Πίνακας 5 παρουσιάζει τον πίνακα σύγχυσης για την κατηγοριοποίηση με βάση τη φασματική γωνία.

ERROR MATRIX [pixel count]									
	> Reference								
V_Classified	1	2	3	4	5	6	7	8	Total
1	91	21	0	40	0	0	0	0	152
2	0	59	0	1	0	0	0	0	60
3	0	0	97	0	0	0	0	0	97
4	0	10	0	7	0	1	12	0	30
5	0	0	0	0	114	0	0	0	114
6	0	0	0	34	0	100	0	0	134
7	0	0	0	0	0	0	81	0	81
8	0	0	0	0	0	0	0	78	78
Total	91	90	97	82	114	101	93	78	746
Overall accuracy [%] = 84.0483									

Πίνακας 5: Πίνακας σύγχυσης φασματικής γωνίας

Στη συνέχεια αναλύονται τα αποτελέσματα, με βάση την ακρίβεια κατηγοριοποίησης παραγωγού, χρήστη αλλά και της συνολικής ακρίβειας.

Ακρίβειες ταξινόμησης παραγωγού:

- 1 (Ελαιώνας): $91/91 \cdot 100\% = 100\%$
- 2 (Συνεχής αστικός ιστός): $59/90 \cdot 100\% = 65.55\%$
- 3 (Νερό): $97/97 \cdot 100\% = 100\%$
- 4 (Κωνοφόρο Δάσος): $7/82 \cdot 100\% = 8.53\%$
- 5 (Ασυνεχής αστικός ιστός): $114/114 \cdot 100\% = 100\%$
- 6 (Καλλιέργειες): $100/101 \cdot 100\% = 99\%$
- 7 (Γυμνό έδαφος): $81/93 \cdot 100\% = 87.09\%$
- 8 (Καμένη περιοχή): $78/78 \cdot 100\% = 100\%$

Στην παραπάνω ανάλυση αποτυπώνεται αναλυτικά πόσα εικονοστοιχεία από κάθε κατηγορία έχουν ταξινομηθεί σωστά. Σε τέσσερις κατηγορίες επιτυγχάνεται απόλυτα ορθή ταξινόμηση, σε δύο κατηγορίες τα ποσοστά επιτυχούς ταξινόμησης ανέρχονται σε ικανοποιητικά επίπεδα, ωστόσο στην κατηγορία 4 τα ποσοστά επιτυχούς ταξινόμησης είναι ιδιαίτερα χαμηλά.

Ακρίβειες ταξινόμησης χρήστη:

- 1 (Ελαιώνας): $91/152 \cdot 100\% = 59.86\%$
- 2 (Συνεχής αστικός ιστός): $59/60 \cdot 100\% = 98.33\%$
- 3 (Νερό): $97/97 \cdot 100\% = 100\%$
- 4 (Κωνοφόρο Δάσος): $7/30 \cdot 100\% = 23.33\%$
- 5 (Ασυνεχής αστικός ιστός): $114/114 \cdot 100\% = 100\%$
- 6 (Καλλιέργειες): $100/134 \cdot 100\% = 74.62\%$
- 7 (Γυμνό έδαφος): $81/81 \cdot 100\% = 100\%$
- 8 (Καμένη περιοχή): $78/78 \cdot 100\% = 100\%$

Και σε αυτή την περίπτωση παρατηρούμε τέσσερις κατηγορίες που εμφανίζουν σφάλματα στις κατηγοριοποιήσεις των εικονοστοιχείων τους, με τα ποσοστά σφαλμάτων να κυμαίνονται σε διάφορα επίπεδα. Η συνολική ακρίβεια στην παρούσα ταξινόμηση είναι 84.04%, δηλαδή παραμένει σε υψηλό επίπεδο. Κατά την εφαρμογή των τριών διαφορετικών μεθόδων ταξινόμησης, παρατηρήθηκαν διαφοροποιήσεις στη συμπεριφορά ορισμένων κατηγοριών κάλυψης γης, καθώς και η επαναλαμβανόμενη παρουσία σφαλμάτων ταξινόμησης σε συγκεκριμένες κατηγορίες, όπως τα κωνοφόρα δάση. Τα σφάλματα αυτά αποδίδονται στην αδυναμία επαρκούς διαχωρισμού των φασματικών υπογραφών μεταξύ ορισμένων κατηγοριών, καθώς και στα στατιστικά χαρακτηριστικά των δεδομένων εκπαίδευσης, τα οποία δεν ήταν επαρκώς αντιπροσωπευτικά. Επιπλέον, η περιορισμένη έκταση και η σχετικά χαμηλή ποικιλομορφία των πολυγώνων εκπαίδευσης συνέβαλαν στην εμφάνιση σφαλμάτων, καθιστώντας δυσχερή την αποτελεσματική εκπαίδευση των αλγορίθμων. Η αξιολόγηση των αποτελεσμάτων πραγματοποιήθηκε μέσω δύο μεθοδολογικών προσεγγίσεων: αφενός της ποιοτικής αξιολόγησης, η οποία βασίστηκε στην οπτική ανάλυση των ταξινομημένων εικόνων που παρήχθησαν μέσω του SCP, και αφετέρου της ποσοτικής αξιολόγησης, η οποία στηρίχθηκε στα αριθμητικά αποτελέσματα που προέκυψαν από τους πίνακες σύγχυσης. Από την ποσοτική ανάλυση προέκυψε ότι η μέθοδος της μέγιστης πιθανοφάνειας εμφάνισε υψηλότερη ακρίβεια συγκριτικά με τις υπόλοιπες μεθόδους. Ωστόσο, το γεγονός ότι η αξιολόγηση στηρίχθηκε αποκλειστικά στα περιορισμένα πολύγωνα εκπαίδευσης υποδηλώνει ότι τα ποσοτικά αποτελέσματα δεν αποτυπώνουν πλήρως την πραγματική ποιότητα της ταξινόμησης. Η έλλειψη επαρκών και διαφοροποιημένων δειγμάτων εκπαίδευσης προκάλεσε αφενός σύγχυση στην αποτύπωση της φασματικής υπογραφής των κατηγοριών και αφετέρου στατιστικά σφάλματα, τα οποία επηρέασαν τη δυνατότητα ακριβούς ταξινόμησης ακόμα και μέσω της μεθόδου της μέγιστης πιθανοφάνειας.

9 Συμπεράσματα

Η παρούσα εργασία ανέδειξε με τρόπο συστηματικό τη δυναμική που προσφέρει η συνδυασμένη χρήση τεχνικών μηχανικής μάθησης και γεωγραφικών συστημάτων πληροφοριών στην ανάλυση και πρόβλεψη γεωχωρικών δεδομένων. Η θεωρητική μελέτη απέδειξε ότι η μηχανική μάθηση παρέχει δυνατότητες πέρα από τις παραδοσιακές τεχνικές GIS, επιτρέποντας την αυτοματοποίηση διαδικασιών, την αποκάλυψη κρυφών προτύπων και τη βελτίωση της ακρίβειας ταξινομήσεων σε πολυδιάστατους γεωγραφικούς χώρους.

Η πειραματική εφαρμογή στην περιοχή της Κορινθίας έδειξε ότι, παρά τους περιορισμούς που επιβάλλουν τα μικρά σύνολα δεδομένων εκπαίδευσης και οι ιδιαιτερότητες της φασματικής συμπεριφοράς κάποιων τύπων κάλυψης γης, οι αλγόριθμοι μηχανικής μάθησης προσφέρουν σαφώς ανώτερη αποδοτικότητα σε σχέση με παραδοσιακές μεθόδους. Η μέθοδος της μέγιστης πιθανοφάνειας αποδείχθηκε αποτελεσματική όσον αφορά την ποσοτική αξιολόγηση, προσφέροντας υψηλά επίπεδα συνολικής ακρίβειας, ιδίως σε περιπτώσεις όπου τα δεδομένα εκπαίδευσης πληρούσαν τις απαιτούμενες στατιστικές προϋποθέσεις κανονικότητας και ομοιογένειας.

Αντίστοιχα, η μέθοδος της φασματικής γωνίας επέδειξε καλύτερη συμπεριφορά σε επίπεδο οπτικής ανάλυσης, καθώς κατάφερε να αποτυπώσει με μεγαλύτερη λεπτομέρεια και αξιοπιστία τη δομή και τη γεωμετρική συνέπεια των χρήσεων γης. Αυτό αναδεικνύει τη σημασία επιλογής της κατάλληλης μεθόδου ταξινόμησης ανάλογα με τον στόχο: σε εφαρμογές όπου το ζητούμενο είναι η χωρική ακρίβεια και η αναπαράσταση της γεωμετρίας των αντικειμένων, οι τεχνικές βασισμένες σε φασματικές ομοιότητες υπερτερούν. Η ανάλυση ανέδειξε, επίσης, ότι η επιτυχία οποιασδήποτε ταξινόμησης βασίζεται σε τρεις κύριες παραμέτρους: την ποιότητα των δεδομένων εκπαίδευσης, την προσεκτική επιλογή των καναλιών που χρησιμοποιούνται στην ανάλυση και την ορθή κατανόηση των ιδιαιτεροτήτων κάθε αλγορίθμου. Η ύπαρξη ανομοιογενών ή ελλιπών δεδομένων εκπαίδευσης οδηγεί σε φαινόμενα σύγχυσης κατηγοριών και σε μείωση της συνολικής ακρίβειας. Η ανεπάρκεια φασματικής διακρίσιμότητας μεταξύ γειτονικών κατηγοριών, όπως η περίπτωση μεταξύ ελαιώνων και γυμνού εδάφους ή μεταξύ διαφορετικών τύπων αστικού ιστού, επηρέασε σημαντικά την ποιότητα των αποτελεσμάτων.

Επιπλέον, αναδείχθηκε η ανάγκη ύπαρξης πολυδιάστατης αξιολόγησης (ποιοτικής και ποσοτικής), καθώς κάθε μέθοδος ταξινόμησης μπορεί να παρουσιάζει διαφορετικά πλεονεκτήματα αναλόγως του τρόπου ανάλυσης των αποτελεσμάτων. Η συνδυασμένη χρήση πινάκων σύγχυσης και ποιοτικής οπτικής εξέτασης αποτελεί αναγκαία διαδικασία για τη σφαιρική αποτίμηση της ταξινόμησης.

Τέλος, διαπιστώθηκε ότι οι δυνατότητες της μηχανικής μάθησης σε συνδυασμό με τα GIS δεν έχουν ακόμη εξαντληθεί. Η εισαγωγή νέων τεχνικών, όπως τα νευρωνικά δίκτυα βαθιάς μάθησης, η ανάλυση υπερφασματικών δεδομένων και η ενσωμάτωση δεδομένων διαφορετικών πηγών (fusion), αναμένεται να ανοίξει νέους δρόμους στην ανάλυση γεωχωρικών δεδομένων. Η παρούσα εργασία, με τη θεωρητική της θεμελίωση και την πρακτική εφαρμογή, θέτει ένα σταθερό υπόβαθρο για μελλοντικές ερευνητικές προσπάθειες στον τομέα αυτό.

10 Επίλογος

Η ταχεία εξέλιξη της τεχνολογίας, σε συνδυασμό με τη δραματική αύξηση του όγκου γεωχωρικών δεδομένων που συλλέγονται καθημερινά από δορυφόρους, αισθητήρες και drones, καθιστά επιτακτική την ανάγκη για νέες, πιο αποτελεσματικές μεθόδους επεξεργασίας και ανάλυσης. Σε αυτό το πλαίσιο, η μηχανική μάθηση δεν λειτουργεί απλώς ως εργαλείο υποστήριξης, αλλά ως βασικός μοχλός μετασχηματισμού της γεωγραφικής επιστήμης και της χαρτογράφησης.

Η εργασία αυτή ανέδειξε με σαφήνεια ότι ο συνδυασμός GIS και μηχανικής μάθησης οδηγεί σε νέα επίπεδα κατανόησης του χώρου, παρέχοντας εργαλεία που είναι πιο ακριβή, πιο αποδοτικά και ικανά να αντεπεξέλθουν στις προκλήσεις της σύγχρονης εποχής. Μέσα από την εφαρμογή αλγορίθμων σε δεδομένα τηλεπισκόπησης, αποδείχθηκε ότι η ανάλυση μεγάλων και πολύπλοκων συνόλων δεδομένων μπορεί να αυτοματοποιηθεί σε μεγάλο βαθμό, μειώνοντας το κόστος, τον απαιτούμενο χρόνο και την πιθανότητα ανθρώπινου σφάλματος.

Παράλληλα, διαφάνηκε η ανάγκη για αυστηρότερα πρότυπα ποιότητας στα δεδομένα εκπαίδευσης και για μεγαλύτερη κατανόηση των περιορισμών κάθε αλγορίθμου. Η τεχνητή νοημοσύνη και η μηχανική μάθηση δεν προσφέρουν «μαγικές λύσεις», αλλά εργαλεία που αποδίδουν στο μέγιστο όταν η διαδικασία σχεδιασμού, συλλογής και προετοιμασίας των δεδομένων είναι επιστημονικά τεκμηριωμένη και μεθοδικά υλοποιημένη.

Η παρούσα εργασία έθεσε τις βάσεις για την εφαρμογή πιο σύνθετων μεθόδων στο μέλλον, όπως η χρήση μετα-αλγορίθμων (ensemble methods), η ανάπτυξη μοντέλων βάθους (deep learning) σε γεωχωρικά δεδομένα και η αξιοποίηση real-time δεδομένων για δυναμική ανάλυση. Σε επόμενα στάδια, η γεωχωρική ανάλυση δεν θα περιορίζεται μόνο στην απεικόνιση και καταγραφή του χώρου, αλλά θα επεκτείνεται στη βαθύτερη κατανόηση της χωροχρονικής δυναμικής και στην προγνωστική ανάλυση.

11 Βιβλιογραφία

Géron, A. (2022). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (3rd ed.). O'Reilly Media.

Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep Learning. MIT Press.

Hackeling, G., 2017 Mastering Machine Learning with scikit-learn. Packt Publishing Ltd.

Kenji Ose, 2018, Introduction to GDAL Tools in QGIS

Kuhlman, D. 2009, A python book: Beginning python, advanced python, and python exercises. Dave Kuhlman Lutz, 2009.

Mukhamediev, Ravil I., 2022, et al. "Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges." Mathematics 10.15 : 2552

Oliphant, T. E., 2006, A guide to NumPy, vol. 1. Trelgol Publishing USA, p. 36.

Shi, Wenzhong, 2022, et al. "Change detection based on artificial intelligence: State-of-the-art and challenges." Remote Sensing 12.10 :1688.

Van Rossum, G., 2007, "Python Programming Language.," in USENIX annual technical conference, vol. 41,

Vozenilek, Vit. 2009, "Artificial intelligence and GIS: mutual meeting and passing." 2009 International conference on intelligent networking and collaborative systems. IEEE.

Βαλδεσέρα Χριστίνα – Λυδία, 2023, Αυτόματη εξαγωγή οροφών κτιρίων με μεθόδους Μηχανικής Μάθησης για την χωροθέτηση ηλιακών φωτοβολταϊκών πάνελ

Δημαράκη Φωτεινή, 2022, Εφαρμογή Τεχνικών Μηχανικής Μάθησης και Εξόρυξης Δεδομένων στην Γεωργία Ακριβείας

Κάβουρας, Μ., Δάρρα, Α., Κονταξάκη, Σ., & Τομαή, Ε. (2015). Επιστήμη Γεωγραφικής Πληροφορίας

Καπαγερίδης Ι., 2006. Εισαγωγή στα Γεωγραφικά Συστήματα Πληροφοριών, Κοζάνη.

Κουτσόπουλος Κ., Ανδρουλακάκης Ν., 2012. Γεωγραφικά Συστήματα Πληροφοριών-Θεωρία & Πράξη. Εκδόσεις Παπασωτηρίου, σελ.641, Αθήνα.

Φιλιππίδης Ε.Ι., 2006. Γεωγραφικά Συστήματα Πληροφοριών. Διδακτικές Σημειώσεις. Τεχνολογικό Εκπαιδευτικό Ίδρυμα Σερρών, Τμήμα Πληροφορικής & Επικοινωνιών.

Χρήστος Κυριάκος, 2022, BUSINESS INTELLIGENCE THROUGH MACHINE LEARNING FROM SATELLITE REMOTE SENSING DATA