



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΒΙΟΪΑΤΡΙΚΗΣ

Τίτλος εκπονούμενης διατριβής:

«Αναζήτηση βιοδραστικών ενώσεων έναντι διαμεμβρανικών υποδοχέων και μελέτη των μηχανισμών δομής – δράσης τους με μεθόδους μοριακής μοντελοποίησης και Μηχανικής Μάθησης»

2^η ΕΚΘΕΣΗ ΠΡΟΟΔΟΥ
1/1/23 – 30/12/23

Ονοματεπώνυμο Υποψήφιου Διδάκτορα: Βασίλειος Παναγιωτόπουλος

Αριθμός Μητρώου: 2103

Επιβλέπων Καθηγητής: Ματσούκας Μίνως-Τιμόθεος, Επίκουρος Καθηγητής

Τριμελής Συμβουλευτική Επιτροπή:

Ματσούκας Μίνως-Τιμόθεος, Επίκουρος Καθηγητής

Καλατζής Ιωάννης, Καθηγητής

Ζουμπουλάκης Παναγιώτης, Αναπληρωτής Καθηγητής

Αναζήτηση βιοδραστικών ενώσεων έναντι διαμεμβρανικών υποδοχέων και μελέτη των μηχανισμών δομής – δράσης τους με μεθόδους μοριακής μοντελοποίησης και Μηχανικής Μάθησης

Μελέτη διαμεμβρανικών υποδοχέων GPCR

Κατά την διάρκεια του δεύτερου έτους της διδακτορικής διατριβής έχει πραγματοποιηθεί ανασκόπηση στην βιβλιογραφία γύρω από μια υποοικογένεια των GPCRs και συγκεκριμένα γύρω από τους υποδοχείς CRF₁R και CRF₂R [1][2][3]. Οι υποδοχείς αυτοί παίζουν σημαντικό ρόλο στην διατήρηση της ομοιόστασης, καθώς ρυθμίζουν ενδοκρινικές και νευρικές λειτουργίες. Η δυσλειτουργία στους συγκεκριμένους υποδοχείς στον ανθρώπινο οργανισμό, σχετίζεται με την εκδήλωση ασθενειών όπως είναι το άγχος και η κατάθλιψη [4]. Έχει πραγματοποιηθεί δομική ανάλυση των δύο αυτών υποδοχέων και έχουν μελετηθεί ως προς την εκλεκτικότητα που παρουσιάζουν στην δέσμευση τόσο ενδογενών πεπτιδίων, όσο και συνθετικών πεπτιδίων [5][6][7]. Για την περαιτέρω κατανόηση της εκλεκτικότητας των συγκεκριμένων υποδοχέων δημιουργήθηκαν και ομόλογα μοντέλα που απεικονίζουν την δέσμευση πεπτιδίων στους υποδοχείς. Στο πλαίσιο αυτής της δομικής ανάλυσης συντάσσεται και μια δημοσίευση ανασκόπησης, η οποία θα κατατεθεί τους προσεχείς μήνες.

Επιπλέον, πραγματοποιήθηκε η αξιολόγηση ενός συνόλου ενώσεων που έχουν συντεθεί από το Department of Immunology, Centre for Neuroendocrinology, University of Pretoria και το εργαστήριο της Επικ. Καθ. Claire Newton με μεθόδους μοριακής πρόσδεσης ως προς την ανενεργή κατάσταση του υποδοχέα GHRHR. Ο GHRHR έχει συσχετιστεί πρόσφατα με αρκετούς τύπους καρκίνου, όπως καρκίνος του προστάτη, καρκίνος του μαστού, καρκίνος στο πάγκρεας και στις ωοθήκες [8][9]. Συγκεκριμένα, πραγματοποιήθηκαν υπολογισμοί μοριακής πρόσδεσης για 3 ενώσεις που παρουσιάζουν ανταγωνιστική δράση ως προς τον συγκεκριμένο υποδοχέα και μελετήθηκαν οι μοριακές αλληλεπιδράσεις των ενώσεων με τα αμινοξέα του ενεργού κέντρου του. Η συγκεκριμένη ερευνητική μελέτη παρουσιάστηκε ως αναρτημένη εργασία (poster) στο 8^ο ERNEST Meeting που πραγματοποιήθηκε στην Κρήτη στις 3-7 Μαΐου 2023 με τίτλο «**V. Panagiotopoulos**, T. Radomsky, R. Perez, K. Tsianakas, M. Giatro, T. Katsila, R. Millar, G. Spyroulias, C. Newton*, M.-T. Matsoukas*. Targeting the Growth Hormone-Releasing Hormone receptor with small molecule antagonists. 8th ERNEST Meeting, 3-7 May 2023, Kolympari, Crete, Greece».

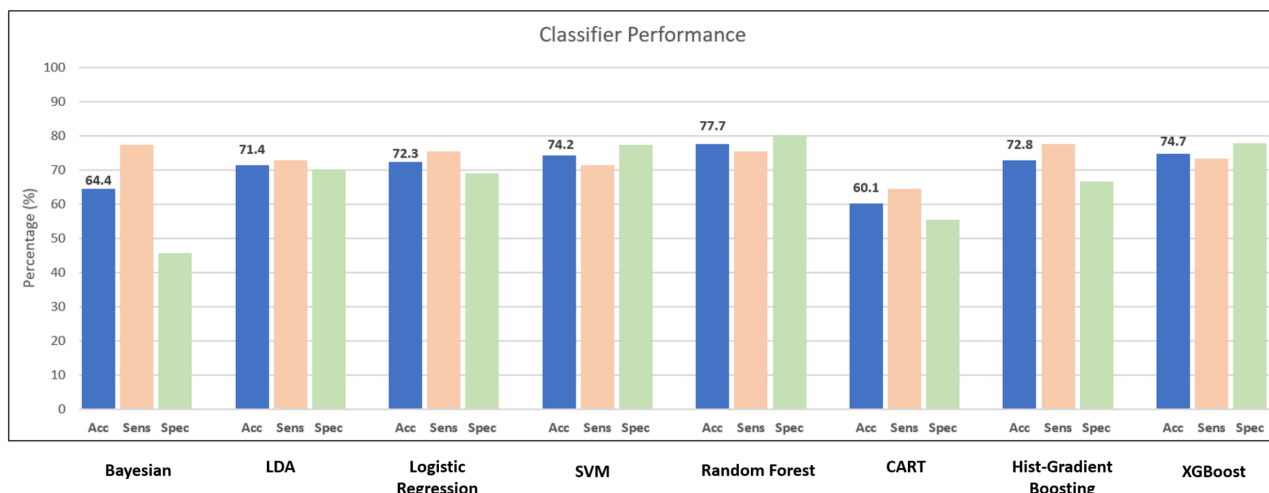
Αλγόριθμοι Μηχανικής Μάθησης στην μοριακή πρόσδεση

Επιπρόσθετα, αναπτύχθηκε μοντέλο μηχανικής μάθησης για αύξηση της αξιοπιστίας της μεθόδου της μοριακής πρόσδεσης. Η υπολογιστική τεχνική της μοριακής πρόσδεσης στηρίζεται σε στοχαστικούς και εμπειρικούς αλγόριθμους, οι οποίοι εμπεριέχουν ένα ποσοστό αβεβαιότητας στις προβλέψεις τους [10][11]. Στόχος ήταν η ανάπτυξη ενός μοντέλου μηχανικής μάθησης το οποίο θα παρείχε ένα ποσοστό βεβαιότητας στις προβλέψεις της μοριακής πρόσδεσης. Ο τρόπος για να επιτευχθεί κάτι τέτοιο ήταν η σύγκριση της προβλεπόμενης διαμόρφωσης ενός μορίου με τη διαμόρφωση ενός πειραματικού μορίου. Για την υλοποίηση λήφθηκαν από την Protein Data Bank [12] περίπου 11.000 πρωτεΐνες, οι οποίες είχαν έναν συγκρυσταλλωμένο προσδέτη. Στην συνέχεια πραγματοποιήθηκαν προσομοιώσεις μοριακής πρόσδεσης του συγκρυσταλλωμένου προσδέτη στην αντίστοιχη πρωτεΐνη

και ως αποτέλεσμα λήφθηκαν 9 προβλεπόμενες διαμορφώσεις (πόζες) του προσδέτη στην κοιλότητα της πρωτεΐνης. Για κάθε ζεύγος πόζας-πρωτεΐνης υπολογίστηκε ένα μοριακό αποτύπωμα αλληλεπιδράσεων μέσω εργαλείων χημειοπληροφορικής. Το μοριακό αποτύπωμα που χρησιμοποιήθηκε έχει σταθερό μήκος 168 bits και κωδικοποιεί το σύνολο των αλληλεπιδράσεων του προσδέτη με τα αμινοξέα της κοιλότητας πρόσδεσης της πρωτεΐνης [13][14][15]. Στην συνέχεια έγινε σύγκριση κάθε προβλεπόμενης διαμόρφωσης που προέκυψε από την μοριακή πρόσδεση με την πειραματική διαμόρφωση χρησιμοποιώντας ως μέτρο ποσοτικοποίησης της σύγκρισης, την τετραγωνική ρίζα σύγκλισης τυπικής απόκλισης (RMSD). Όσο μεγαλύτερη είναι η τιμή του RMSD τόσο πιο διαφορετική διαμόρφωση έχει η προβλεπόμενη πόζα σε σχέση με την πειραματική. Για τιμές RMSD μικρότερες ή ίσες με 2 Angströms, θεωρείται ότι οι διαμορφώσεις δεν παρουσιάζουν μεγάλη διαφορά μεταξύ τους και η προβλεπόμενη πόζα μοιάζει με την πειραματική. Μια πρόβλεψη που πληροί τα παραπάνω χαρακτηριστικά θεωρείται ‘καλή’ πρόβλεψη της μοριακής πρόσδεσης (ευνοϊκή πόζα). Στόχος για την ανάπτυξη του μοντέλου μηχανικής μάθησης ήταν η δυνατότητα να μπορεί να διακρίνει τις ‘καλές’ διαμορφώσεις (ευνοϊκές πόζες) που προκύπτουν από τους υπολογισμούς μοριακής πρόσδεσης από τις ‘κακές’ (μη ευνοϊκές πόζες). Μέτρο για την ανάπτυξη του μοντέλου αυτού αποτέλεσε η απόκλιση των συντεταγμένων της πόζας των προσδέτων ως προς αυτές της κρυσταλλικής δομής του πειραματικού μορίου. Επομένως, η τιμή του RMSD για την κάθε προβλεπόμενη πόζα χρησιμοποιήθηκε για την κατηγοριοποίηση (labeling) των δεδομένων. Οι προβλεπόμενες διαμορφώσεις, στις οποίες το RMSD σε σχέση με την κρυσταλλική δομή ήταν μικρότερο ή ίσο με 2 Angströms κατηγοριοποιήθηκαν στην ‘Θετική Κλάση’, ενώ οι προβλεπόμενες διαμορφώσεις στις οποίες το RMSD σε σχέση με την κρυσταλλική δομή ήταν μεγαλύτερο ή ίσο με 4 Angströms και μικρότερο ή ίσο με 6 Angströms ($4\text{\AA} \leq \text{RMSD} \leq 6\text{\AA}$) κατηγοριοποιήθηκαν στην ‘Αρνητική Κλάση’. Στην συνέχεια το σύνολο δεδομένων (dataset) χωρίστηκε σε training set και test set με αναλογία 70-30% και το training set χρησιμοποιήθηκε για την εκπαίδευση και την επικύρωση 8 ταξινομητών μηχανικής μάθησης. Επειδή υπήρχε ανομοιομορφία στο σύνολο των δεδομένων ανάμεσα στις δύο κλάσεις (class imbalance), εφαρμόστηκε η μέθοδος random undersampling ώστε οι δύο κλάσεις να περιέχουν ίσο αριθμό δειγμάτων. Με την μέθοδο αυτή, τα δείγματα της μεγαλύτερης κλάσης μειώνονται με τυχαίο τρόπο ώστε να φτάσουν τον αριθμό δειγμάτων της μικρότερης κλάσης. Επιπλέον, αφαιρέθηκαν τα χαρακτηριστικά τα οποία είχαν χαμηλή διακύμανση και μεγάλη συσχέτιση μεταξύ τους.

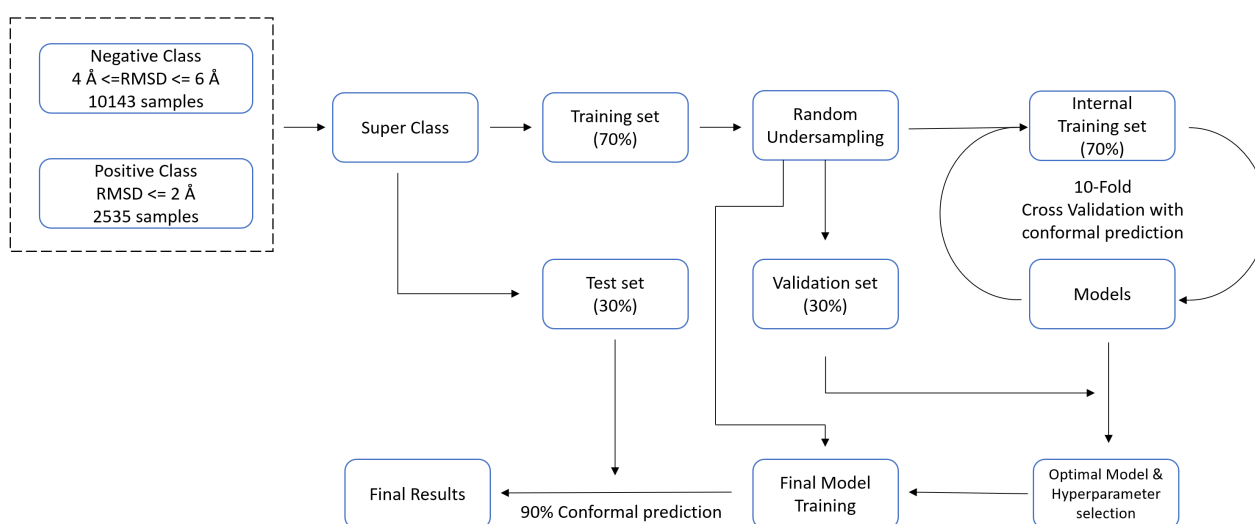
Στην συνέχεια το training set χωρίστηκε εκ νέου σε ένα καινούριο training set, το οποίο ονομάστηκε ως ‘internal training set’ για να μην υπάρχει σύγχυση με το αρχικό training set και ένα validation set πάλι σε αναλογία 70-30%. Πραγματοποιήθηκε η εκπαίδευση και η επικύρωση των ταξινομητών χρησιμοποιώντας την μέθοδο k-Fold cross validation για 10 folds στο internal training set και το validation set χρησιμοποιήθηκε ως external test set για κάθε ταξινομητή. Στην πορεία εξήχθησαν οι τιμές της ακρίβειας (accuracy), ειδικότητας (specificity) και ευαισθησίας (sensitivity) για κάθε ταξινομητή για το validation set. Για κάθε ταξινομητή έγινε και hyperparameter tuning, ώστε να βρεθούν οι καλύτερες παράμετροι από ένα εύρος τιμών χρησιμοποιώντας την μέθοδο GridSearch. Οι επιδόσεις των ταξινομητών φαίνονται στην **Εικόνα 1**.

Αναζήτηση βιοδραστικών ενώσεων έναντι διαμεμβρανικών υποδοχέων και μελέτη των μηχανισμών δομής – δράσης τους με μεθόδους μοριακής μοντελοποίησης και Μηχανικής Μάθησης



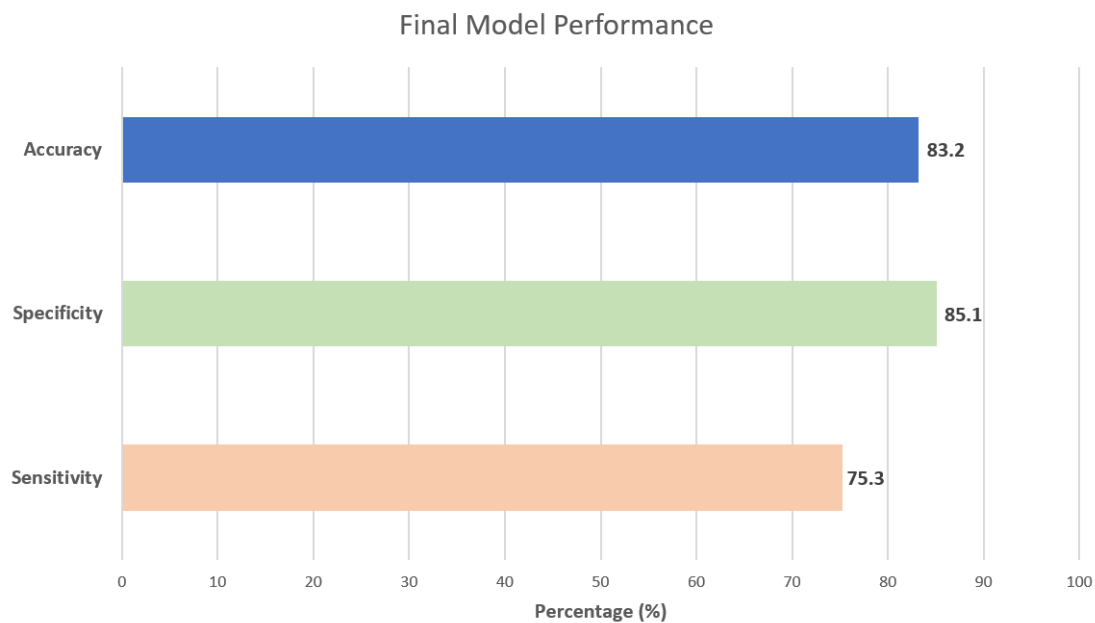
Εικόνα 1. Αποτελέσματα ταξινόμητών κατά την επικύρωση με το validation set.

Στην συνέχεια επιλέχθηκε ο ταξινόμητής ο οποίος είχε την καλύτερη επίδοση, ο Random Forest, και εκπαιδεύτηκε εκ νέου σε όλο το αρχικό training set (70% όλων των δεδομένων). Η επικύρωση του ταξινόμητή έγινε με το αρχικό test set το οποίο δεν συμμετείχε σε κανένα στάδιο της εκπαίδευσης. Επιπλέον, για την επικύρωση των δεδομένων χρησιμοποιήθηκε και το πλαίσιο της σύμμορφης πρόβλεψης (conformal prediction), με το οποίο εξασφαλίζεται μια επιπλέον αξιοπιστία στις προβλέψεις του ταξινόμητή [17][18]. Το επιθυμητό επίπεδο εμπιστοσύνης ορίστηκε στο 90%, δηλαδή λήφθηκαν υπόψιν τα δείγματα στα οποία η ακρίβεια των προβλέψεων του ταξινόμητή ήταν μεγαλύτερη από 90% για κάθε κλάση. Η ακρίβεια του ταξινόμητή ήταν 83.2%, η ειδικότητά του ήταν 85.1% και η ευαισθησία του ήταν 75.3%. Στην **Εικόνα 2** φαίνεται ένα σχεδιάγραμμα της μεθοδολογίας, στην **Εικόνα 3** φαίνονται οι επιδόσεις του βέλτιστου ταξινόμητή και στην **Εικόνα 4** φαίνεται η καμπύλη ROC του ταξινόμητή (ROC Curve).

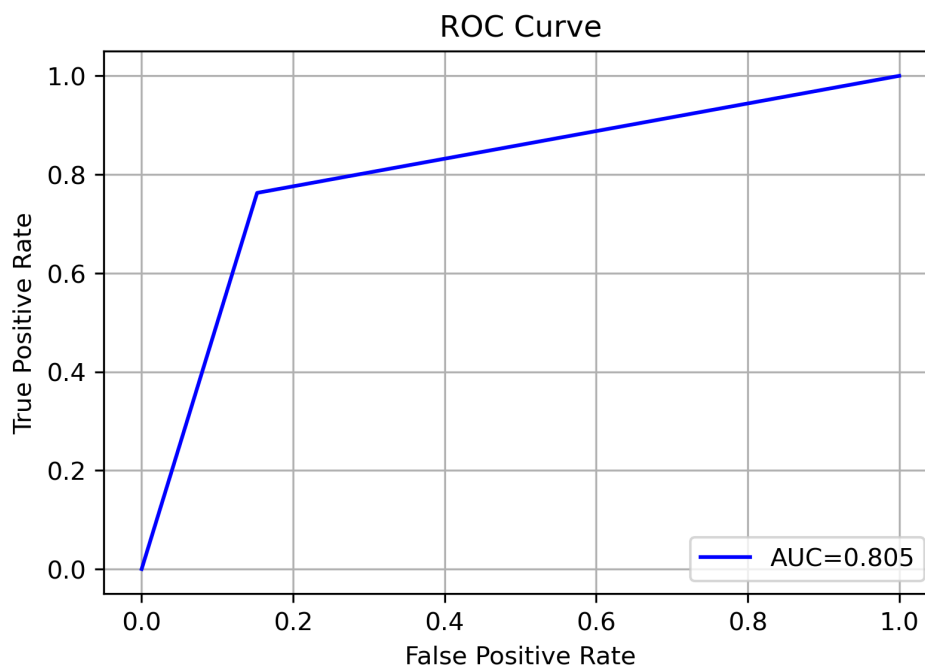


Αναζήτηση βιοδραστικών ενώσεων έναντι διαμεμβρανικών υποδοχέων και μελέτη των μηχανισμών δομής – δράσης τους με μεθόδους μοριακής μοντελοποίησης και Μηχανικής Μάθησης

Εικόνα 2. Σχεδιάγραμμα μεθοδολογίας



Εικόνα 3. Αποτελέσματα βέλτιστου ταξινομητή στο σύνολο των δεδομένων.



Εικόνα 4. Καμπύλη ROC του βέλτιστου ταξινομητή για το σύνολο των δεδομένων.

Αναζήτηση βιοδραστικών ενώσεων έναντι διαμεμβρανικών υποδοχέων και μελέτη των μηχανισμών δομής – δράσης τους με μεθόδους μοριακής μοντελοποίησης και Μηχανικής Μάθησης

Η συγκεκριμένη ερευνητική μελέτη παρουσιάστηκε ως αναρτημένη εργασία (poster) στο συνέδριο Advances in Biomedical Sciences, Engineering and Technology (ABSET) που πραγματοποιήθηκε στην Αθήνα στις 10-11 Ιουνίου 2023 με τίτλο «**V. Panagiotopoulos**, S. Ouzounis, M. Giatro, P. Zoumpoulakis, D. Cavouras, I. Kalatzis, T. Katsila, M.-T. Matsoukas*. Structure-Based Computational Drug Repurposing using Interaction Fingerprint Data: A Machine Learning Approach. Advances in Biomedical Sciences, Engineering and Technology (ABSET), 10-11 June 2023, Athens, Greece». Καθώς επίσης παρουσιάστηκε ως αναρτημένη εργασία (poster) και στο συνέδριο Biomedicine, Bioinformatics, Biotechnology Forum το οποίο έλαβε χώρα στην Αθήνα στις 15-17 Σεπτεμβρίου 2023 με τίτλο «**V. Panagiotopoulos**, S. Ouzounis, K. Tsianakas, M. Giatro, P. Zoumpoulakis, T. Katsila, M.-T. Matsoukas*. A Machine Learning Approach for Structure-Based Computational Drug Repurposing Leveraging Interaction Fingerprint Data. Biomedicine, Biotechnology, Bioinformatics Forum, 15-17 September 2023, Athens, Greece».

Αλγόριθμοι Μηχανικής Μάθησης για τον Δάγκειο πυρετό

Ο Δάγκειος πυρετός είναι μια μόλυνση η οποία οφείλεται στον δάγκειο ιό και μεταδίδεται μέσω των κουνουπιών. Υπάρχουν 4 διαφορετικοί σερότυποι του Δάγκειου ιού που απαντώνται ως DENV-1, DENV-2, DENV-3 και DENV-4. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας τις δύο τελευταίες δεκαετίες οι περιπτώσεις εμφάνισης του Δάγκειου πυρετού έχουν αυξηθεί κατά 10 φορές με τον αριθμό να ξεπερνάει τα 5.2 εκατομμύρια περιστατικών το 2019. Μέχρι σήμερα δεν υπάρχουν διαθέσιμα φάρμακα στην αγορά για την αντιμετώπιση του Δάγκειου πυρετού, οπότε είναι σημαντικό να βρεθούν νέες ενώσεις, οι οποίες να είναι δραστικές έναντι των σερότυπων του Δάγκειου ιού. Σκοπός της παρούσας ερευνητικής εργασίας είναι η δημιουργία μοντέλων Μηχανικής Μάθησης, τα οποία θα μπορούν να προβλέψουν αν μια χημική ένωση είναι δραστική έναντι στον Δάγκειο ιό [18][19][20].

Τα μοντέλα αυτά θα εκπαιδευτούν με δεδομένα χημειοπληροφορικής όπως είναι οι μοριακοί περιγραφείς (molecular descriptors) και τα μοριακά αποτυπώματα (molecular fingerprints) [21]. Για να γίνει αυτό έπρεπε να αντληθεί πληροφορία για τις δραστικές και μη δραστικές ενώσεις έναντι του Δάγκειου ιού που είναι εργαστηριακά επιβεβαιωμένες. Η πληροφορία αυτή λήφθηκε από την δημόσια βάση δεδομένων ChEMBL [22] όπου για κάθε έναν από τους 4 σερότυπους του Δάγκειου ιού αντλήθηκαν οι ενώσεις για τις οποίες είχε μετρηθεί πειραματικά η τιμή EC50, η οποία εκφράζει την δραστικότητα μιας ένωσης σε μία πρωτεΐνη. Τα δεδομένα που αντλήθηκαν ήταν ο ChEMBL κωδικός της κάθε ένωσης, τα SMILES, το μοριακό βάρος και η τιμή EC50 (μετρημένη σε nM) και δημιουργήθηκαν 4 datasets, ένα για κάθε σερότυπο του ιού. Τα SMILES είναι μία ASCII συμβολοσειρά η οποία περιγράφει την δομή της χημικής ένωσης. Στην συνέχεια έγινε προεπεξεργασία των δεδομένων, όπου κρατήθηκαν ενώσεις οι οποίες είχαν μοριακό βάρος μεγαλύτερο από 200 Dalton και μικρότερο από 800 Dalton, αφαιρέθηκαν τυχόν άλατα από τα SMILES και αφαιρέθηκαν τυχόν διπλότυπες ενώσεις. Ακολούθως, έγινε η εξαγωγή των μοριακών περιγραφέων για κάθε ένωση χρησιμοποιώντας το πακέτο λογισμικού RDKit [23] της python. Για κάθε ένωση υπολογίστηκαν 208 μοριακοί περιγραφείς, οι οποίοι εκφράζουν φυσικοχημικές ιδιότητες και χαρακτηριστικά της ένωσης, όπως το

Αναζήτηση βιοδραστικών ενώσεων έναντι διαμεμβρανικών υποδοχέων και μελέτη των μηχανισμών δομής – δράσης τους με μεθόδους μοριακής μοντελοποίησης και Μηχανικής Μάθησης

μοριακό βάρος, η λιποφιλικότητα, ο αριθμός των δακτυλίων κλπ. Την παρούσα στιγμή πραγματοποιούνται δοκιμές στην εκπαίδευση αλγορίθμων Μηχανικής Μάθησης, ώστε να βρεθούν οι βέλτιστες παράμετροι για την διαχωρίσιμότητα των δεδομένων.

Μελλοντικοί στόχοι της διδακτορικής διατριβής είναι:

- η εικονική σάρωση μεγάλων βιβλιοθηκών ενώσεων, όπως είναι η Enamine REAL (~6 εκατ. ενώσεις) [24], έναντι του υποδοχέα GHRHR με στόχο τον εντοπισμό βιοδραστικών ενώσεων.
- Η βελτιστοποίηση του μοντέλου Μηχανικής Μάθησης για αύξηση της αξιοπιστίας της μεθοδολογίας της μοριακής πρόσδεσης (αύξηση δεδομένων, αλλαγή μοριακού αποτυπώματος, χρήση δικτύων Βαθιάς Μάθησης)
- Η ανάπτυξη μοντέλων Μηχανικής Μάθησης για την πρόβλεψη της δραστηριότητας των ενώσεων έναντι του Δάγκειου ιού χρησιμοποιώντας διαφορετικού τύπου μοριακούς περιγραφείς και μοριακών αποτυπωμάτων.
- Η ανάπτυξη μοντέλων Μηχανικής Μάθησης και Βαθιάς Μάθησης για την πρόβλεψη της δραστηριότητας μικρών ενώσεων σε πρωτεΐνες του ανθρώπινου γονιδιώματος, αξιοποιώντας δομική πληροφορία χαρακτηρισμένων πρωτεϊνών είτε ομόλογων μοντέλων υψηλής ακρίβειας από τις προβλέψεις του AlphaFold [25].

Τέλος, παρατίθενται επιστημονικές δημοσιεύσεις στις οποίες υπήρχε συμμετοχή:

- Ouzounis, S., **Panagiotopoulos, V.**, Bafiti, V., Zoumpoulakis, P., Cavouras, D., Kalatzis, I., Matsoukas, M. T., & Katsila, T. (2023). A Robust Machine Learning Framework Built Upon Molecular Representations Predicts CYP450 Inhibition: Toward Precision in Drug Repurposing. *Omics : a journal of integrative biology*, 27(7), 305–314. <https://doi.org/10.1089/omi.2023.0075>
- Ouzounis, S., Kanterakis, A., **Panagiotopoulos, V.**, Cavouras, D., Zoumpoulakis, P., Matsoukas, M. T., Katsila, T., Kalatzis, I. (2023). Data-Driven Drug Repurposing in Diabetes Melitus through an Enhanced Knowledge Graph. *Engineering Proceedings*, 50(1), 9. <https://doi.org/10.3390/engproc2023050009>

Βιβλιογραφία

- [1] Wootten D, Christopoulos A, Marti-Solano M, Babu MM, Sexton PM. Mechanisms of signalling and biased agonism in G protein-coupled receptors. *Nature Reviews Molecular Cell Biology* (2018) 19: 638
- [2] Sriram K, Insel PA. G protein-coupled receptors as targets for approved drugs: How many targets and how many drugs?. *Molecular Pharmacology* (2018) 93: 251
- [3] Yang D, Zhou Q, Labroska V, Qin A, Darbalaei S, Wu Y, Yuliantie E, Xie L, Tao H, Cheng J, Liu Q, Zhao S, Shui W, Jiang Y, Wang MW. G protein-coupled receptors: structure- and function-based drug discovery. *Signal Transduction and Targeted Therapy* (2021) 6: 7
- [4] Ma S, Shen Q, Zhao LH, Mao C, Zhou XE, Shen DD, de Waal PW, Bi P, Li C, Jiang Y, Wang MW, Sexton PM, Wootten D, Melcher K, Zhang Y, Xu HE. Molecular Basis for Hormone Recognition and Activation of Corticotropin-Releasing Factor Receptors. *Molecular Cell* (2020) 77: 669
- [5] Liang YL, Belousoff MJ, Zhao P, Koole C, Fletcher MM, Truong TT, Julita V, Christopoulos G, Xu HE, Zhang Y, Khoshouei M, Christopoulos A, Danev R, Sexton PM, Wootten D. Toward a Structural Understanding of Class B GPCR Peptide Binding and Activation. *Molecular Cell* (2020) 77: 656
- [6] Islam MR, Teleb M, Karageorgos V, Sakellaris S, Papadopoulos M, Pirmettis I, Fronczek FR, Liapakis G, Fahmy H. Design, synthesis, structural optimization, SAR, in silico prediction of physicochemical properties and pharmacological evaluation of novel & potent thiazolo[4,5-d]pyrimidine corticotropin releasing factor (CRF) receptor antagonists. *European journal of pharmaceutical sciences* (2022) 169: 106084.
- [7] Cordomi A, Liapakis G, Matsoukas M-T. Understanding Corticotropin Releasing Factor Receptor (CRFR) Activation Using Structural Models. *Current Molecular Pharmacology* (2017) 10:325
- [8] Zhou F, Zhang H, Cong Z, Zhao LH, Zhou Q, Mao C, Cheng X, Shen DD, Cai X, Ma C, Wang Y, Dai A, Zhou Y, Sun W, Zhao F, Zhao S, Jiang H, Jiang Y, Yang D, Eric Xu H, Zhang Y, Wang MW. Structural basis for activation of the growth hormone-releasing hormone receptor. *Nature Communications* (2020) 11: 5205
- [9] Matsoukas M-T, Spyroulias GA. Dynamic properties of the growth hormone releasing hormone receptor (GHRHR) and molecular determinants of GHRH binding. *Molecular BioSystems* (2017) 13: 1313
- [10] Pinzi L, Rastelli G. Molecular Docking: Shifting Paradigms in Drug Discovery. *International Journal of Molecular Sciences* (2019) 20:4331
- [11] Saikia S, Bordoloi M. Molecular Docking: Challenges, Advances and its Use in Drug Discovery Perspective. *Current drug targets* (2019) 20:501
- [12] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic acids research* (2000) 28:253
- [13] Bouysset C, Fiorucci S. ProLIF: a library to encode molecular interactions as fingerprints. *Journal of cheminformatics* (2021) 13:72
- [14] Wang DD, Chan MT, Yan H. Structure-based protein-ligand interaction fingerprints for binding affinity prediction. *Computational and structural biotechnology journal* (2021) 19:6291

Αναζήτηση βιοδραστικών ενώσεων έναντι διαμεμβρανικών υποδοχέων και μελέτη των μηχανισμών δομής – δράσης τους με μεθόδους μοριακής μοντελοποίησης και Μηχανικής Μάθησης

- [15] Vass M, Kooistra AJ, Ritschel T, Leurs R, de Esch IJ, de Graaf C. Molecular interaction fingerprint approaches for GPCR drug discovery. *Current opinion in pharmacology* (2016) 30:59
- [16] Alvarsson J, Arvidsson McShane S, Norinder U, Spjuth O. Predicting With Confidence: Using Conformal Prediction in Drug Discovery. *Journal of pharmaceutical sciences* (2021) 110:42
- [17] Krstajic D. Critical Assessment of Conformal Prediction Methods Applied in Binary Classification Settings. *Journal of chemical information and modeling* (2021) 61:4823
- [18] Khew C, Akbar R, Mohd-Assaad N. Progress and challenges for the application of machine learning for neglected tropical diseases. *F1000Research* (2023) 12:287
- [19] S AH, Pujar GV, Sethu AK, Bhagyalalitha M, Singh M. Dengue structural proteins as antiviral drug targets: Current status in the drug discovery & development. *European journal of medicinal chemistry* (2021) 221:113527
- [20] Lim SP. Dengue drug discovery: Progress, challenges and outlook. *Antiviral research* (2019) 163:156
- [21] Lo YC, Rensi SE, Torng W, Altman RB. Machine learning in chemoinformatics and drug discovery. *Drug Discovery Today* (2018) 23: 1538
- [22] Davies M, Nowotka M, Papadatos G, Dedman N, Gaulton A, Atkinson F, Bellis L, Overington JP. ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* (2015) 43(W1): W612
- [23] RDKit: Open-source cheminformatics. <https://www.rdkit.org>
- [24] Enamine REAL compounds (2020). Available at: <https://enamine.net/library-synthesis/real-compounds>
- [25] Jumper J et al. Highly accurate protein structure prediction with AlphaFold. *Nature* (2021) 596: 583

Ο επιβλέπων:

Ματσούκας Μίνως-
Τιμόθεος

**Ο υποψήφιος
διδάκτορας:**

Παναγιωτόπουλος
Βασίλειος



Επίκουρος Καθηγητής