

Εθνικό Μετσόβιο Πολυτεχνείο



Μεταπτυχιακή Διπλωματική Εργασία

Υλοποίηση συναισθηματικά επαυξημένου
υβριδικού νευρωνικού δικτύου

Συγγραφέας:
Ανθούσα Καρκόγλου

Επιβλέπων:
Ανδρέας-Γεώργιος
Σταφυλοπάτης

Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών
Μηχανικής Μάθησης και Επιστήμης Δεδομένων
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ

8 Απριλίου 2021

Εθνικό Μετσόβιο Πολυτεχνείο



Μεταπτυχιακή Διπλωματική Εργασία

Υλοποίηση συναισθηματικά επαυξημένου
υβριδικού νευρωνικού δικτύου

Συγγραφέας:
Ανθούσα Καρκόγλου

Επιβλέπων:
Ανδρέας-Γεώργιος
Σταφυλοπάτης

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή στις 8 Απριλίου 2021

Ανδρέας-Γεώργιος
Σταφυλοπάτης

Γεώργιος Στάμου

Κωνσταντίνα Νικήτα

Καθηγητής Ε.Μ.Π.

Αναπληρωτής
Καθηγητής Ε.Μ.Π.

Καθηγήτρια Ε.Μ.Π.

Copyright ©– All rights reserved Ανθούσα Καρχόγλου, 2021

Με επιφύλαξη παντός δικαιώματος. **All rights reserved.**

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Συναισθηματική υπολογιστική συνιστά ο επιστημονικός τομέας της ενσωμάτωσης του συναισθήματος σε υπολογιστικά συστήματα με εγγενή τρόπο. Η εμφάνιση του όρου πραγματοποιήθηκε στα μέσα της τελευταίας δεκαετίας του 20^{ου} αιώνα και βασίζεται σε μελέτες που υπογραμμίζουν τη σημασία του συναισθήματος στην ανθρώπινη νόηση. Ο τομέας αυτός απαντάται σε πολλές εφαρμογές που, είτε έχουν στόχο να εμφυσήσουν συναίσθημα σε μοντέλα, έτσι ώστε να επιδεικνύουν συναισθηματική συμπεριφορά, είτε να προσομοιώνουν εγκεφαλικές δομές που υλοποιούν την ροή της συναισθηματικής πληροφορίας και άλλων νοητικών λειτουργιών. Ακόμα συμπεριλαμβάνονται εφαρμογές που αναγνωρίζουν το συναίσθημα στην αλληλεπίδρασή τους με τον χρήστη.

Στόχος της διπλωματικής εργασίας είναι η σύντομη παρουσίαση του θεωρητικού υπόβαθρου, αλλά και των σημαντικότερων εφαρμογών του τομέα αυτού σε μοντέλα μηχανικής μάθησης. Ανάμεσα σε αυτά τα μοντέλα επιλέγεται ένα υβριδικό δίκτυο το οποίο υλοποιείται και στη συνέχεια αξιολογείται στα πλαίσια ταξινόμησης σε δεδομένα αναγνώρισης συναισθήματος.

Λέξεις κλειδιά

συναισθηματική υπολογιστική, τεχνητή νοημοσύνη, υβριδικό νευρωνικό δίκτυο, χάρτες αυτο-οργάνωσης (SOM), συναισθηματική οπισθοδιάδοση, μεταιχμιακό σύστημα, αναγνώριση προσώπου, αναγνώριση συναισθήματος

Abstract

Affective computing addresses the integration of sentiment into computational systems in an intrinsic way. This scientific area was introduced during the last decade of the 20th century and is founded upon theories that underline the importance of sentiment in human cognition. Affective computing can be found in applications that aim at embedding affective behavior into computational systems or simulating anatomical areas of the brain that regulate emotion among other cognitive functions. Other applications within this area address the recognition of emotion during the interaction between human and machine.

The scope of this thesis includes a concise presentation of the theoretical background of affective computing alongside the description of a number of significant emotion-augmented machine learning models. Amongst them, a hybrid neural network is selected and implemented in order to be evaluated against an emotion recognition task.

Keywords

affective computing, artificial intelligence, hybrid network, self-organizing maps (SOM), emotional backpropagation, limbic system, facial recognition, sentiment analysis, emotion recognition

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον καθηγητή Ανδρέα-Γεώργιο Σταφυλοπάτη για την ευκαιρία που μου δόθηκε να διερευνήσω ένα ιδιαίτερα ενδιαφέρον θέμα στον χώρο επικάλυψης των επιστημονικών πεδίων της τεχνητής νοημοσύνης και της συναισθηματικής υπολογιστικής. Ακόμα οφείλω να ευχαριστήσω τον Γιώργο Σιόλα για την άψογη συνεργασία μας, την υπομονή που επέδειξε και την καθοδήγησή του καθ'όλη τη διάρκεια εκπόνησης της παρούσας μεταπτυχιακής διπλωματικής εργασίας.

Περιεχόμενα

Περίληψη	v
Abstract	vii
Ευχαριστίες	ix
1 Εισαγωγή	1
1.1 Έναυσμα	1
1.2 Αντικείμενο της Διπλωματικής	1
1.3 Δομή της Διπλωματικής	2
2 Φυσιολογία του συναισθήματος	3
2.1 Συναίσθημα	3
2.1.1 Μοντελοποίηση συναισθήματος	3
2.1.2 Ανατομία συναισθήματος	4
2.1.3 Συναίσθημα στην λήψη αποφάσεων	8
2.2 Συναισθηματική υπολογιστική	9
2.2.1 Εγκαινίαση ερευνητικού τομέα	9
2.2.2 Ανάπτυξη και πρακτικές εφαρμογές	10
3 Συναισθηματικά επαυξημένη τεχνητή νοημοσύνη (Emotion-augmented AI)	13
3.1 Συναίσθημα σε διεργασίες βελτιστοποίησης	13
3.2 Ανατομικά μοντέλα στην τεχνητή νοημοσύνη	15
3.3 Συναίσθημα στην υπολογιστική ενισχυτική μάθηση	16
3.4 Νόηση και αφαιρετικότητα στη μάθηση	17
4 Υβριδικό δίκτυο συναισθηματικής μάθησης	19
4.1 Χάρτες αυτο-οργάνωσης (Self-Organizing Maps, SOM)	19
4.1.1 Γενικά	19
4.1.2 Δομή	19
4.1.3 Μάθηση	20
4.1.4 Σύγκλιση	21
4.1.5 Εφαρμογές	21
4.2 Συναισθηματικά επαυξημένο SOM (EmSOM)	22
4.2.1 Θεωρητικό υπόβαθρο	22
4.2.2 Δομή νευρωνικού δικτύου EmSOM	24
4.2.2.1 Κρυφό επίπεδο νευρώνων του EmSOM	25
4.2.2.2 Επίπεδο νευρώνων εξόδου του EmSOM	27
4.2.3 Συναισθηματική οπισθοδιάδοση (EmBP)	28
4.2.3.1 Υπολογισμός συναισθηματικών παραμέτρων	28
4.2.3.2 Ανανέωση παραμέτρων δικτύου κατά την εκπαίδευση	29

5	Υλοποίηση και αποτελέσματα	31
5.1	Επιλογή εργαλείων υλοποίησης	31
5.2	Υλοποίηση	31
5.2.1	Σύνολα Δεδομένων	32
5.2.1.1	Προ-επεξεργασία δεδομένων	32
5.2.2	Βασική Δομή Δικτύου	33
5.2.3	Υλοποίηση δικτύου SOM	33
5.2.4	Υλοποίηση του Emotional Backpropagation (EmBP)	34
5.3	Αποτελέσματα	35
5.3.1	Εφαρμογή στο ORL σύνολο δεδομένων	35
5.3.2	Εφαρμογή στο KDEF σύνολο δεδομένων	36
6	Συμπεράσματα	39
6.1	Συμπεράσματα υλοποίησης	39
6.2	Μελλοντικές επεκτάσεις	40
	Βιβλιογραφία	41

Κατάλογος εικόνων

2.1	Μεταχαιμιακό σύστημα (limbic system) [11]	4
2.2	Ροή πληροφορίας στο μεταχαιμιακό σύστημα [11]	7
4.1	Τοπολογία SOM	20
4.2	Είσοδος/Έξοδος του κρυμμένου νευρώνα στο δίκτυο εφαρμογής του EmBP	22
4.3	Είσοδος/Έξοδος του νευρώνα εξόδου στο δίκτυο εφαρμογής του EmBP	23
4.4	EmSOM τοπολογία	24
4.5	Ροή πληροφορίας για το κρυφό επίπεδο νευρώνων.	26
4.6	Ροή πληροφορίας για το επίπεδο νευρώνων εξόδου.	27
5.1	Εικόνες από το ORL σύνολο δεδομένων	32
5.2	Εικόνες από το KDEF σύνολο δεδομένων	32
5.3	Διαγράμματα ορθότητας (accuracy) εκπαίδευσης και αξιολόγησης στο ORL dataset	36
5.4	Ορθότητα (accuracy) εκπαίδευσης και αξιολόγησης ως προς τις εποχές εκπαίδευσης για το KDEF σύνολο δεδομένων	37

Κατάλογος πινάκων

5.1	Ποσοστά ορθότητας (accuracy) εκπαίδευσης για διαφορετικές τιμές εποχών και μεγέθους test δεδομένων	35
5.2	Ποσοστά ορθότητας (accuracy) αξιολόγησης για διαφορετικές τιμές εποχών και μεγέθους test δεδομένων	35
5.3	Ποσοστά ορθότητας για διαφορετικές τιμές εποχών	37

Κεφάλαιο 1

Εισαγωγή

Πριν την παρουσίαση του θεωρητικού υπόβαθρου, της βιβλιογραφίας αλλά και των τεχνικών λεπτομερειών υλοποίησης του επιλεγμένου νευρωνικού δικτύου, παρατίθενται στο κεφάλαιο αυτό το έναυσμα, ο σκοπός αλλά και η δομή της παρούσας διπλωματικής εργασίας.

1.1 Έναυσμα

Η επιστήμη της βιολογίας έχει αποτελέσει έμπνευση της τεχνητής νοημοσύνης από την απαρχή της. Από την δημιουργία του τεχνητού νευρώνα για τη δόμηση σύνθετων νευρωνικών δικτύων που προσομοιώνουν το κεντρικό νευρικό σύστημα και την έμπνευση της οπισθοδιάδοσης (**backpropagation**) από την θεωρία του Freud (Κάθεξις) [56] μέχρι την δημιουργία συνελικτικών νευρωνικών δικτύων βάσει της έρευνας του οπτικού φλοιού αιλουροειδών [57], η τεχνητή νοημοσύνη στοχεύει στην προσομοίωση βιολογικών δομών και συμπεριφορών.

Συνεπώς, θεωρείται λογικό επόμενο ότι με την εμφάνιση επιστημονικών ερευνών που αποδίδουν στο συναίσθημα καταλυτικό ρόλο στην ανθρώπινη νόηση, τις αντίστοιχες εγκεφαλικές λειτουργίες και ιδιαίτερα στην λήψη αποφάσεων θα εγκαινιαζόταν ένα νέο επιστημονικό πεδίο στην επιστήμη των υπολογιστών με τον όρο συναισθηματική υπολογιστική (**affective computing**). Στην ομώνυμη δημοσίευση [38] η συναισθηματική υπολογιστική παρουσιάζεται ως ένας νέος τομέας έρευνας, που περιλαμβάνει ανάμεσα σε άλλα την αναγνώριση ανθρώπινου συναίσθηματος μέσω εκφράσεων και φωνής ως ένα νέο δυναμικό πρόβλημα αναγνώρισης προτύπων. Ακόμα, παρουσιάζονται πιθανές εφαρμογές της συναισθηματικής υπολογιστικής σε ετερόκλητους τομείς όσον αφορά την αναγνώριση συναίσθηματος, ενώ εισάγεται η ιδέα της δημιουργίας υπολογιστών που πιθανά να επιδεικνύουν "συναισθηματικά χαρακτηριστικά" στην κατεύθυνση της ευφυούς συμπεριφοράς και της σωστής λήψης αποφάσεων.

1.2 Αντικείμενο της Διπλωματικής

Στόχος της παρούσας διπλωματικής εργασίας είναι η διερεύνηση της βιβλιογραφίας όσον αφορά την συμπερίληψη συναίσθηματος στην υλοποίηση νευρωνικών δικτύων και εφαρμογών τεχνητής νοημοσύνης. Από αυτή τη βιβλιογραφική έρευνα επιλέγεται ένα νευρωνικό δίκτυο το οποίο υλοποιείται με τη βοήθεια του **Pytorch framework** και εφαρμόζεται σε διαφορετικά δεδομένα με σκοπό την αξιολόγησή του. Το επιλεγμένο νευρωνικό δίκτυο είναι υβριδικό, συνδυάζοντας χάρτες αυτο-οργάνωσης (**Self-Organizing Maps, SOM**) και ένα προστιοδρομικό δίκτυο (**feed-forward network**), ενώ περιλαμβάνει την συναισθηματική παράμετρο τόσο στη δομή του, όσο και στον αλγόριθμο μάθησης που χρησιμοποιεί [9]. Το εν λόγω δίκτυο εφαρμόζεται σε δεδομένα εικόνων προς αναγνώριση προσώπου και στη συνέχεια διερευνάται η απόδοσή του σε δεδομένα εικόνων προσώπου από διαφορετικό σύνολο δεδομένων με σκοπό την αναγνώριση συναίσθηματος.

1.3 Δομή της Διπλωματικής

Παρακάτω παρουσιάζονται τα κεφάλαια της διπλωματικής εργασίας και μία σύντομη περιγραφή του περιεχομένου τους:

- Κεφάλαιο 1: Εισαγωγή και συνοπτική παρουσίαση του σκοπού της διπλωματικής εργασίας.
- Κεφάλαιο 2: Παρουσίαση βασικών εννοιών και θεωριών όσον αφορά το συναίσθημα και τη φυσιολογία του. Ακόμα, στο κεφάλαιο αυτό περιλαμβάνεται μία περιληπτική περιγραφή της ερευνητικής εργασίας που εισήγαγε την συναισθηματική υπολογιστική (**affective computing**), καθώς και η εξέλιξη αυτού του τομέα με ορισμένες εφαρμογές.
- Κεφάλαιο 3: Περιγραφή της βιβλιογραφίας για τα συναισθηματικά νευρωνικά δίκτυα (**emotional neural networks**) ανά κατηγορία προσέγγισης συναισθηματικής υλοποίησης.
- Κεφάλαιο 4: Περιγραφή και ανάλυση του επιλεγμένου υβριδικού συναισθηματικού δικτύου, του **EmSOM**, καθώς και θεωρητική συνοπτική παρουσίαση του βασικού δομικού στοιχείου του, του χάρτη αυτο-οργάνωσης (**Self-Organizing Maps, SOM**).
- Κεφάλαιο 5: Ανάλυση της υλοποίησης του επιλεγμένου δικτύου στο **PyTorch framework** και παρουσίαση των αποτελεσμάτων της εφαρμογής του σε δύο διαφορετικά σύνολα δεδομένων.
- Κεφάλαιο 6: Παρουσίαση συμπερασμάτων και πιθανών μελλοντικών επεκτάσεων.

Κεφάλαιο 2

Φυσιολογία του συναισθήματος

Ο όρος συναισθηματική υπολογιστική εισήχθη από την R.W. Picard το 1995 με κίνητρο την δημιουργία μηχανών με συναισθηματική νοημοσύνη, συμπεριλαμβανομένης και της ενσυναίσθησης. Αποτελεί διεπιστημονικό τομέα που βασίζεται στην επιστήμη των υπολογιστών, την ψυχολογία και την γνωσιακή επιστήμη. Στο παρόν κεφάλαιο περιλαμβάνεται μία συνοπτική θεώρηση διαφορετικών πτυχών της θεωρίας του συναισθήματος και ακολουθεί συνοπτική παρουσίαση της συναισθηματικής υπολογιστικής.

2.1 Συναίσθημα

Εδώ παρουσιάζονται συνοπτικά βασικές έννοιες και θεωρίες που αποτελούν θεμέλια για την ανάπτυξη των νευρωνικών δικτύων που παρουσιάζονται στο Κεφάλαιο 3. Λόγω των διαφορετικών προσεγγίσεων προσομοίωσης των συναισθημάτων από κάθε νευρωνικό δίκτυο, η περιγραφή αυτή περιλαμβάνει όλες τις θεωρητικές πληροφορίες που συμβάλουν στην κατανόησή τους.

2.1.1 Μοντελοποίηση συναισθήματος

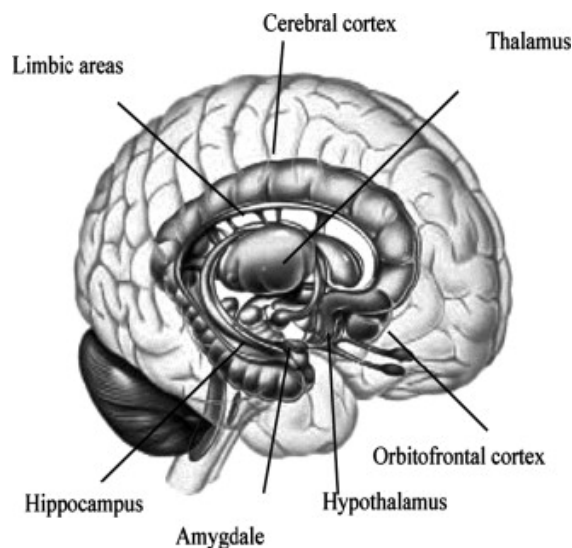
Τα συναισθήματα είναι γνωσιακές διεργασίες που μελετώνται από διάφορες επιστήμες όπως η ψυχολογία, η νευροεπιστήμη και η τεχνητή νοημοσύνη. Από ψυχολογική σκοπιά τα συναισθήματα μπορούν να αντληθούν από επιβράβευση και τιμωρία υπό διάφορες συνθήκες στην πραγματική ζωή. Οι συναισθηματικές αναπαραστάσεις λειτουργούν ως ψυχολογικά κατασκευάσματα για μοντελοποίηση, ανάλυση και αναγνώριση συναισθήματος, ενώ οι πιο σημαντικές μοντελοποιήσεις προκύπτουν από κατηγορικές και βάσει αξιολόγησης προσεγγίσεις, όπως και προσεγγίσεις αναπαραστάσης μέσω διαστάσεων.

Η κατηγορική θεωρία συναισθήματος εκλαμβάνει το σύστημα των συναισθημάτων ως ένα περιορισμένο σύνολο από διακριτές οντότητες, όπως τα βασικά συναισθήματα. Ένα πολύ διαδεδομένο μοντέλο αποτελεί το 'Big 6' του Paul Ekman που διακρίνει κάθε συναίσθημα στις κατηγορίες του θυμού, της αηδίας, του φόβου, της χαράς, της λύπης και της έκπληξης [48]. Στον αντίποδα αυτού, οι υποστηρικτές της αναπαραστάσης συναισθημάτων σε διαστάσεις μοντελοποιούν το συναίσθημα ως ένα πολυδιάστατο χωρικό συνεχές που προσδιορίζεται από τον άξονα της πολικότητας (valence) και της έντασης (arousal) σύμφωνα με το μοντέλο κυκλικής συνθετότητας (circumplex model) [16]. Ωστόσο το πιο συχνά απαντημένο μοντέλο πολυδιάστατης αναπαραστάσης συναισθημάτων περιλαμβάνει τρεις διαστάσεις, την ένταση (arousal), την πολικότητα (valence) και την επικράτηση (dominance) [48]. Το πρότυπο αξιολόγησης (appraisal paradigm) σχηματίζει τη θεμελιώδη βάση για διάφορα υπολογιστικά μοντέλα και ιδιαίτερα μοντέλα ενισχυτικής μάθησης. Σύμφωνα με την θεωρία αξιολόγησης (appraisal theory), τα συναισθήματα προκύπτουν από τις αξιολογήσεις (appraisals, estimates) των γεγονότων που προκαλούν διαφορετικές αντιδράσεις. Ουσιαστικά, η συναισθηματική κατάσταση προκύπτει σύμφωνα με την αξιολόγηση των συνθηκών που επικρατούν.

Επιπρόσθετα, το συναίσθημα έχει γίνει αντιληπτό ως ένας μηχανισμός επιβράβευσης-τιμωρίας για τις εγκεφαλικές λειτουργίες. Σε καθημία από τις μελέτες που υποστηρίζουν αυτή την προσέγγιση το συναίσθημα περιγράφεται να δρα ως ένα σύστημα καθοδήγησης με επικεντρωμένη προσοχή σε συγκεκριμένα σημεία σε ένα περίπλοκο περιβάλλον και ζυγίζοντας την σημασία τους ανάλογα. Ουσιαστικά, το συναίσθημα μπορεί να ταξινομηθεί σε δύο κλάσεις, τα κύρια και τα δευτερεύοντα συναισθήματα. Τα κύρια συναισθήματα αποτελούν εγγενείς αντιδράσεις σε περιβαλλοντικά ερεθίσματα και εστιάζουν στις ανάγκες ενός συστήματος, όπως για παράδειγμα ο συσχετισμός της μυρωδιάς του φαγητού με την ανάγκη κατανάλωσης φαγητού. Τα δευτερεύοντα συναισθήματα προκύπτουν από τον συσχετισμό τους με κύρια ερεθίσματα και, επομένως, με συσχετιστικά συναισθήματα που μαθαίνονται για παράδειγμα μέσω ανάδρασης. Αυτές οι συσχετίσεις καθίστονται επίκτητοι παράγοντες πρόβλεψης για τις περιπτώσεις που προκύπτει μία κύρια συναισθηματική αντίδραση [16].

2.1.2 Ανατομία συναισθήματος

Η νευρολογική βάση του συναισθήματος βρίσκεται στο μεταιχμιακό σύστημα (limbic system) του εγκεφάλου. Το σύστημα αυτό αποτελείται από τις κύριες δομές της αμυγδαλής (amygdala), του κορχομετωπιαίου φλοιού (orbitofrontal cortex, OFC), του θαλάμου (thalamus), του αισθητικού φλοιού (sensory cortex), του υποθαλάμου (hypothalamus), και του ιππόκαμпу (hippocampus). Στην συνέχεια παρουσιάζονται πιο αναλυτικά οι επιμέρους δομές.



Εικόνα 2.1: Μεταιχμιακό σύστημα (limbic system) [11]

- **Αμυγδαλή (amygdala):** Το όνομα της δομής αυτής προκύπτει από το αμυγδαλοειδές της σχήμα. Τοποθετημένες δίπλα στον ιππόκαμπο, η αριστερή και η δεξιά αμυγδαλή παίζουν κεντρικό ρόλο στις συναισθηματικές αντιδράσεις, συμπεριλαμβανομένων των αισθημάτων της ευχαρίστησης, του άγχους, του φόβου και του θυμού. Η αμυγδαλή προσάπτει συναισθηματικό περιεχόμενο στις μνήμες και για το λόγο αυτό παίζει σημαντικό ρόλο στον καθορισμό της έντασης με την οποία αποθηκεύονται. Οι μνήμες που διαθέτουν έντονο συναισθηματικό νόημα τείνουν να διατηρούνται.

Η αμυγδαλή δεν τροποποιεί απλά την ένταση και το συναισθηματικό περιεχόμενο των μνήμων. Κατέχει καθοριστικό ρόλο στην δημιουργία νέων που σχετίζονται με το φόβο. Μνήμες που πλαισιώνονται από φόβο αποχρυσταλλοποιούνται μετά από

ελάχιστες επαναλήψεις. Για το λόγο αυτό η προσέγγιση της μάθησης μέσω φόβου συνιστά έναν δημοφιλή τρόπο έρευνας των μηχανισμών καθορισμού και ανάκλησης μνήμης [22].

- **Κογχομετωπιαίος φλοιός (orbitofrontal cortex, OFC):** Ο κογχομετωπιαίος φλοιός είναι η περιοχή του μετωπιαίου φλοιού που βρίσκεται πάνω από τις κόγχες των ματιών. Διαθέτει εκτεταμένες συνδέσεις με τις αισθητικές περιοχές του μεταιχμιακού συστήματος που σχετίζονται με συναίσθημα και μνήμη.

Ο κογχομετωπιαίος φλοιός συχνά συσχετίζεται με συγκεκριμένους τύπους λήψης αποφάσεων. Για παράδειγμα, έχει διατυπωθεί η υπόθεση ότι ο κογχομετωπιαίος φλοιός είναι σημαντικός για τις αποφάσεις που πραγματοποιούνται με σύγκριση σχετικών αξιών ανάμεσα σε διάφορες επιλογές, ούτως ώστε να προκύψει η καλύτερη [55]. Ο κογχομετωπιαίος φλοιός συνδέεται με την αμυγδαλή και έχει πραγματοποιηθεί η υπόθεση ότι εμπλέκεται σε σωματικές ρυθμίσεις βάσει του συναίσθηματος. Η υπόθεση αυτή υποστηρίζεται από πειράματα σε ασθενείς με βλάβη στην περιοχή αυτή, όπως το πείραμα με ασκήσεις στοιχημάτων όπου οι συμμετέχοντες με βλάβη στον κογχομετωπιαίο φλοιό έτειναν να πραγματοποιούν πιο ριψοκίνδυνες επιλογές χωρίς σημάδια άγχους [1]. Ο πληθυσμός ελέγχου (control group) πραγματοποίησε στο ίδιο πείραμα λιγότερο ριψοκίνδυνες επιλογές, ενώ στην περίπτωση επιλογής μίας ριψοκίνδυνης απόφασης οι ίδιοι εμφάνιζαν μεγαλύτερη αγωγιμότητα δέρματος που συνιστά μέτρο εφίδρωσης και κατ'επέκταση άγχους. Οι συμμετέχοντες με βλάβη κογχομετωπιαίου φλοιού μπορούσαν να αναγνωρίσουν την ριψοκινδυνότητα των επιλογών τους, αλλά παρόλα αυτά συνέχιζαν να τις πραγματοποιούν. Οι ερευνητές κατέληξαν ότι αυτό το τμήμα του εγκεφάλου εμπλέκεται στην παροχή σημαντικών σωματικών σημάτων που βοηθούν στην αναγνώριση μίας κακής επιλογής μέσω της συναισθηματικής αντίδρασης σε αυτό [1]. Μετέπειτα έρευνες υποστηρίζουν ότι η αδυναμία στις αποφάσεις των ασθενών με βλάβη στον κογχομετωπιαίο φλοιό μπορεί να μην αντιστοιχεί αποκλειστικά σε συναισθηματική παράμετρο, αφού οι ίδιοι δεν παρουσιάζουν ολική έκλειψη συναισθημάτων [53].

Ολόκληρη η εικόνα του ρόλου του κογχομετωπιαίου φλοιού δεν έχει ακόμα κατανοηθεί από την επιστημονική κοινότητα και υπάρχει μεγάλη αβεβαιότητα για τον ακριβή προσδιορισμό των λειτουργιών που συντονίζει. Είναι σίγουρο ωστόσο ότι παίζει σημαντικό ρόλο στην νόηση και το συναίσθημα.

- **Θάλαμος (thalamus):** Ο θάλαμος αποτελείται από δύο λοβούς φαιάς ουσίας (grey matter) που βρίσκονται κάτω από τον εγκεφαλικό φλοιό. Αποτελεί ένα κεντρικό επεξεργαστικό κέντρο για την αισθητική πληροφορία, αφού συνδέει τα συσχετιζόμενα μέρη του εγκεφαλικού φλοιού με την σπονδυλική στήλη και άλλα μέρη του εγκεφάλου που είναι σημαντικά για τις αισθήσεις. Ακόμα, ο θάλαμος ελέγχει τον ύπνο [22].
- **Αισθητικός φλοιός (sensory cortex):** Τα ερεθίσματα που ανιχνεύονται στους αισθητικούς υποδοχείς μεταφράζονται σε δυναμικό ενέργειας, δηλαδή έναν ηλεκτρικό παλμό που μεταδίδεται κατά μήκος ενός προσαγωγού νευρώνα. Τελικά το σήμα αυτό φτάνει σε ένα συγκεκριμένο τμήμα του εγκεφάλου όπου πραγματοποιείται η επεξεργασία και η ερμηνεία της αισθητηριακής πληροφορίας. Τα τμήματα αυτά αποτελούν τον αισθητικό φλοιό, ο οποίος διακρίνεται σε τμήματα κάθε αίσθησης και συνήθως αναφέρονται με τους όρους σωματισταθητικός φλοιός, οπτικός φλοιός, ακουστικός, οσφρητικός και γευστικός φλοιός. Όπως είναι φανερό, καθεμία από αυτές τις δομές αντιστοιχεί σε μία από τις πέντε αισθήσεις. Κάθε αίσθηση έχει δυο πανομοιότυπα "κέντρα", ένα σε κάθε ημισφαίριο του εγκεφάλου. Ο νωτιαίος μυελός παίζει πολύ

σημαντικό ρόλο στη μεταφορά των πληροφοριών, καθώς από αυτόν περνούν όλα τα ερεθίσματα από κάθε σημείο του σώματος πλην του κεφαλιού. Όλες οι πληροφορίες που έρχονται από τα διάφορα αισθητήρια συστήματα, με εξαίρεση το οσφρητικό σύστημα, περνούν πρώτα από τον θάλαμο του εγκεφάλου, ο οποίος στη συνέχεια διανέμει τις πληροφορίες στα αντίστοιχα κέντρα [59].

- **Υποθάλαμος (hypothalamus):** Ο υποθάλαμος αποτελεί ένα μικρό μέρος του εγκεφάλου που βρίσκεται ακριβώς κάτω από τον θάλαμο και στις δύο πλευρές της τρίτης κοιλίας (ventricle). Οι κοιλίες είναι περιοχές του εγκεφάλου (cerebrum) που περιέχουν εγκεφαλονωτιαίο υγρό (cerebrospinal fluid) και το ενώνουν με την σπονδυλική στήλη. Ο υποθάλαμος βρίσκεται στις δύο ζώνες του οπτικού νεύρου και μόλις πάνω από την υπόφυση (pituitary gland) με την οποία και συνδέεται.

Ο υποθάλαμος είναι από τα πιο δραστήρια μέρη του εγκεφάλου και αναλαμβάνει την ομοιοστάση. Η ομοιοστάση είναι μία διαδικασία θερμοστατική, όπου η θερμοκρασία διατηρείται σε συγκεκριμένα επίπεδα ανεξάρτητα των συνθηκών. Ακόμα, ο υποθάλαμος είναι υπεύθυνος μεταξύ άλλων για τη ρύθμιση του αισθήματος της πείνας, της δίψας, της απόκρισης στον πόνο, των επιπέδων ευχαρίστησης και σεξουαλικής ικανοποίησης, του θυμού και της επιθετικής συμπεριφοράς. Επιπρόσθετα ρυθμίζει την λειτουργία του αυτόνομου νευρικού συστήματος, που με τη σειρά του ρυθμίζει τον καρδιακό παλμό, την πίεση του αίματος, την αναπνοή και την διέγερση σε απάντηση συναισθηματικών συνθηκών [3].

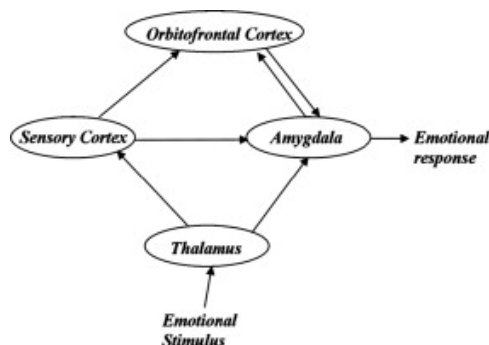
- **Ιππόκαμπος (hippocampus):** Όπως και η αμυγδαλή και άλλες εγκεφαλικές δομές, ο ιππόκαμπος συνιστά ζεύγος δομών καθεμία από τις οποίες βρίσκεται σε διαφορετικό ημισφαίριο. Προσομοιάζει την μορφή ενός θαλάσσιου ιππόκαμπου και αποτελεί ουσιαστικά το κέντρο μνήμης του εγκεφάλου. Εδώ σχηματίζονται οι επεισοδικές μνήμες και συντάσσονται ώστε να αποθηκευτούν ως μεγάλης διάρκειας μνήμη σε άλλα μέρη του φλοιού.

Οι συνδέσεις που δημιουργούνται στον ιππόκαμπο βοηθούν στις συσχετιστικές μνήμες με διάφορες αισθήσεις (όπως για παράδειγμα η μνήμη που ανακαλεί μία γνώριμη μυρωδιά). Ο ιππόκαμπος είναι ακόμα σημαντικός για τον χωρικό προσανατολισμό και την ικανότητα πλοήγησης στον κόσμο.

Ο ιππόκαμπος είναι μία δομή όπου νέοι νευρώνες δημιουργούνται από βλαστοκύτταρα. Αυτή η διαδικασία ονομάζεται νευρογένεση και είναι η βάση της πλαστικότητας του εγκεφάλου. Επομένως είναι αναμενόμενο ότι αυτή η δομή παίζει καταλυτικό ρόλο στην μάθηση νέων πραγμάτων [22].

- **Προμετωπιαίος φλοιός (prefrontal cortex, PFC):** Συνιστά το μέρος του εγκεφαλικού φλοιού που καλύπτει το μπροστινό μέρος του μετωπιαίου λοβού. Η περιοχή αυτή συνδέεται με διατελεστικές διεργασίες όπως τον σχεδιασμό, την λήψη αποφάσεων, τη βραχυπρόθεσμη μνήμη, την έκφραση της προσωπικότητας, όπως και τη ρύθμιση της κοινωνικής συμπεριφοράς, αλλά και ορισμένες ρυθμίσεις της ομιλίας και της γλώσσας. Η βασική δραστηριότητα αυτής της περιοχής θεωρείται ο συντονισμός σχέσεων και δράσεων σε ευθυγράμμιση με εσωτερικούς στόχους [58]. Μέρος του προμετωπιαίου φλοιού συνιστά ο κογχομετωπιαίος φλοιός.

Ο ρόλος των επιμέρους δομών που περιγράφηκαν παραπάνω δεν μπορεί να διαχωριστεί απόλυτα. Για το λόγο αυτό το συναίσθημα στον εγκέφαλο είναι πολύπλευρο και οι συναισθηματικές λεπτομέρειες προκύπτουν από συνυφασμένες δραστηριότητες μεταξύ των διάφορων συστημάτων [16].



Εικόνα 2.2: Ροή πληροφορίας στο μεταιχμιακό σύστημα [11]

Στην Εικόνα 2.2 απεικονίζεται η σχηματική αλληλεπίδραση της αμυγδαλής με άλλα εγκεφαλικά συστήματα. Η αναπαράσταση αυτή βασίζεται στα συμπεράσματα που προκύπτουν από την έρευνα του LeDoux [30]. Βάσει αυτής ανάμεσα στα μονοπάτια δρομολόγησης σήματος υπάρχουν συντομότερα μονοπάτια μέσω των οποίων τα συναισθηματικά ερεθίσματα επεξεργάζονται πολύ πιο γρήγορα. Πιο συγκεκριμένα, υπάρχουν δύο διαφορετικοί τρόποι για ένα εξωτερικό ερέθισμα να φτάσει στην αμυγδαλή. Ο πρώτος είναι σύντομος και γρήγορος, αλλά χωρίς έμφραση στις λεπτομέρειες και προέρχεται από τον θάλαμο. Ο δεύτερος είναι μακρύς και αργός, αλλά με μεγάλο επίπεδο λεπτομέρειας και προέρχεται από το αισθητικό φλοιό. Η αμυγδαλή είναι με τέτοιο τρόπο τοποθετημένη ώστε να λαμβάνει ένα ερέθισμα υπερβολικά γρήγορα και να προκαλεί την απαραίτητη αντίδραση. Επομένως, ένα συναισθηματικό ερέθισμα όπως ο φόβος μπορεί να προκαλέσει γρήγορες αντιδράσεις όταν δεν υπάρχει πιθανότητα για το λογικό κομμάτι του εγκεφάλου να επεξεργαστεί τον κίνδυνο. Όπως αναφέρθηκε και παραπάνω, η αμυγδαλή συνιστά την αποθήκη συναισθηματικών αναμνήσεων και είναι υπεύθυνη για τα συναισθηματικά ερεθίσματα. Η αμυγδαλή λαμβάνει σήματα επιβράβευσης κατά τη διαδικασία μάθησης και αλληλεπιδρά με τον κορχομετωπιαίο φλοιό. Ο τελευταίος λαμβάνει συνδέσεις από τον αισθητικό φλοιό και την αμυγδαλή. Η αμυγδαλή με τη σειρά της απαντά στο συναισθηματικό ερέθισμα. Στη συνέχεια η απάντηση της αμυγδαλής αξιολογείται από τον κορχομετωπιαίο φλοιό που επιχειρεί να αποτρέψει ακατάλληλες απαντήσεις βάσει των πληροφοριών γενικότερου πλαισίου συνθηκών που παρέχονται από τον ιππόκαμπο [4].

Τέλος, θα πρέπει να σημειωθεί η συνεισφορά του D. O. Hebb στην κατανόηση του τρόπου μάθησης σε γνωσιακό και νευρολογικό επίπεδο. Οι τρεις βασικές παραδοχές στη θεωρία του Hebb έθεσαν τη βάση για την βαθιά κατανόηση της ανθρώπινης συμπεριφοράς σε συνάρτηση με το νευρικό σύστημα. Οι παραδοχές αυτές παρουσιάζονται εδώ συνοπτικά ως βάση όχι μόνο ενός σημαντικού μέρους της συναισθηματικής υπολογιστικής, η οποία θα παρουσιαστεί στη συνέχεια, αλλά και ολόκληρου του επιστημονικού κλάδου της τεχνητής νοημοσύνης [47].

1. Μάθηση του Hebb (Hebbian learning): Σύμφωνα με αυτή την παραδοχή οι συνδέσεις ανάμεσα σε νευρώνες αποκτούν αυξανόμενη αποτελεσματικότητα σε αναλογία με τον βαθμό της συσχέτισης ανάμεσα στην προσυναπτική (pro-synaptic) και μετα-συναπτική (meta-synaptic) δραστηριότητα. Η σύναψη δηλαδή των νευρώνων που βρίσκονται κοντά και συχνά ενεργοποιούνται ταυτόχρονα αποκτά σύνδεση μεγαλύτερης δύναμης ("neurons that fire together wire together").
2. Δομές νευρώνων (cell assemblies): Τη νευρωνική βάση της νοητικής αναπαράστασης ιδεών, εικόνων και άλλων εννοιών αποτελούν οι ομάδες ή συστάδες νευρώνων που τείνουν να ενεργοποιούνται ταυτόχρονα βάσει του τρόπου μάθησης του Hebb (Hebbian learning). Η ενεργοποίηση των νευρώνων σε μία τέτοια δομή μπορεί να

διατηρηθεί μετά από την πυροδότηση και μέσω αυτής της διατήρησης δημιουργείται μίας μορφής μνήμη.

3. Διαδοχή φάσης (**phase sequence**): Αξιοποιώντας τις δύο προηγούμενες έννοιες, η σκέψη ορίζεται ως διαδοχική ενεργοποίηση δομών νευρώνων. Η διαδικασία της σκέψης ορίζεται ως αλληλουχία φάσης (**phase sequence**) και επί της ουσίας αναπαριστά την ενεργοποίηση δομών νευρώνων που μετά από την συχνή ενεργοποίηση και συσταδοποίησή τους αναλαμβάνουν δράση ως κλειστά συστήματα που επικοινωνούν με άλλες τέτοιες δομές και υλοποιούν συγκεκριμένες διεργασίες.

Από τα παραπάνω είναι φανερή η σημασία και ο ρόλος της νευρωνικής πλαστικότητας σε όλες τις γνωσιακές ανθρώπινες λειτουργίες, αλλά κυρίως στην μάθηση μέσω της τροποποίησης της τοπολογίας και της έντασης μεταξύ των συνάψεων νευρώνων.

2.1.3 Συναίσθημα στην λήψη αποφάσεων

Η δημοφιλής κλίμακα **Myers-Briggs** για την αναγνώριση προσωπικότητας τοποθετεί την σκέψη και την αίσθηση ως αντίθετα άκρα ενός από τους άξονες περιγραφής της. Είναι γνωστό ότι υπερβολικά πολλά συναισθήματα μπορούν να προκαλέσουν χάος. Ωστόσο μελέτες δείχνουν ότι υπάρχει μία μετατόπιση στην θεωρία του πώς οι άνθρωποι ξεχωρίζουν τα συναισθήματα και τη λογική [38].

Εδώ αναφέρονται κάποια συμπεράσματα του **Damasio** [10], ο οποίος εξέτασε ασθενείς με βλάβες στον μετωπιαίο λοβό. Οι βλάβες αυτές επηρεάζουν την περιοχή του φλοιού που επικοινωνεί με το μεταιχμιακό σύστημα, ενώ οι ασθενείς παρουσιάζονται ως ευφυείς και ασυνήθιστα λογικοί. Παρόλα αυτά οι ίδιοι εμφανίζουν προβληματική ικανότητα στη λήψη αποφάσεων. Πιο συγκεκριμένα διαπιστώνεται ότι δαπανούν πολύ χρόνο προσπαθώντας να λάβουν μία απόφαση που υπο φυσιολογικές συνθήκες θα μπορούσε να ληφθεί πολύ γρήγορα. Για παράδειγμα, η επιλογή μίας ημερομηνίας για τον προγραμματισμό ενός ραντεβού μπορεί να τους οδηγήσει μέσω υπερβολικά μεγάλων αλυσιδωτών αποφάσεων στην αδυναμία λήψης πραγματικής απόφασης μέχρι μία ημερομηνία να τους επιβληθεί από εξωτερικό παράγοντα που επιλέγει να μην περιμένει πλέον την απάντησή τους.

Ο τραυματισμός του μετωπιαίου λοβού σε αυτούς τους ασθενείς εμπλέκεται στην ικανότητά τους να συνδυάσουν συναισθηματικές μεταιχμιακές απαντήσεις με την κατά τα άλλα φλοιική ικανότητα λήψης αποφάσεων. Η υπόθεση του **Damasio** είναι ότι το συναίσθημα κατέχει έναν ρόλο μεροληψίας (**bias**) στην διαδικασία λήψης αποφάσεων με το να εδραίωνει αξίες που χρησιμοποιούνται στην αξιολόγηση πιθανών εκβάσεων και ουσιαστικά σταματούν μία ατέρμονη λογική αναζήτηση.

Τα ευρήματα του **Damasio** υποστηρίζουν ανεξάρτητα επιστημονικά επιχειρήματα για τον ουσιαστικό ρόλο του συναισθήματος. Οι **Jonson-Laird et al.** επιβεβαιώνουν στην γνωσιακή κοινότητα την ανικανότητα της λογικής να καθορίσει ποια από τις άπειρες εναλλακτικές συμπερασμάτων είναι λογικό να επιλεγεί βάσει ενός συστήματος παραδοχών. Πράγματι, σε περιορισμένο χρονικό πλαίσιο δεν υπάρχει η δυνατότητα αναλογισμού κάθε πιθανού λογικού περιορισμού και σχετικού μονοπατιού. Το συναίσθημα δεν κατέχει απλά το ρόλο επίλυσης διλημάτων σε σίγουρες αποφάσεις, αλλά φαίνεται να είναι απαραίτητο και στην μάθηση αναγκαίων θέσεων για την κατασκευή λογικών αποχρίσεων. Ακόμα, τα ευρήματα αυτά παρέχουν νευρολογική υποστήριξη στο ότι δεν υπάρχει αμιγής λόγος στο υγιές ανθρώπινο εγκέφαλο, αλλά τα συναισθήματα είναι απαραίτητα για υγιή, λογική σκέψη και συμπεριφορά. Οι ασθενείς του **Damasio** είναι αφύσικα λογικοί, παρόμοια με τα συστήματα κανόνων (**rule-based**) που συνιστούν τα παραδοσιακά μοντέλα λήψης αποφάσεων [10].

Βάσει όλων των παραπάνω η **Picard** καταλήγει στο ότι η επιστημονική μελέτη του συναισθήματος δεν θα έπρεπε να αποτελεί απλά μία παράπλευρη μελέτη της γνωσιακής επιστήμης, αφού κατέχει μεγάλη σημασία προς την κατεύθυνση κατανόησης της ανθρώπινης

νοήσης. Οι εμπλοκές αυτής της μελέτης είναι επίσης σημαντικές για την επιστήμη των υπολογιστών και τη βιομηχανία, αφού προκειμένου οι υπολογιστές να είναι πραγματικά αποδοτικοί στην λήψη αποφάσεων θα πρέπει να διαθέτουν μηχανισμούς προσομοίωσης συναισθήματος σε συγχρονισμό με συστήματα βασισμένα σε κανόνες (rule-based) [38].

2.2 Συναισθηματική υπολογιστική

Η συναισθηματική υπολογιστική αποτελεί έναν διεπιστημονικό ερευνητικό τομέα που βασίζεται μεταξύ άλλων σε θεωρίες κοινωνιολογίας, ψυχολογίας, επιστήμης υπολογιστών, γλωσσολογίας και μαθηματικών. Εδώ παρουσιάζεται η απαρχή αυτού του τομέα, η συνοπτική περιγραφή της πορείας του μέχρι σήμερα και μερικών πρακτικών εφαρμογών.

2.2.1 Εγκαινίαση ερευνητικού τομέα

Στην ερευνητική δημοσίευση που η R. W. Picard εισάγει τον όρο συναισθηματική υπολογιστική παρουσιάζεται η σημασία του συναισθήματος στην νόηση και τη λήψη των αποφάσεων και υπογραμμίζεται ο ρόλος του συναισθήματος στη μάθηση. Ακόμα, παρουσιάζονται τα βασικά προβλήματα στην ανάπτυξη υπολογιστών με ικανότητα να αναγνωρίσουν συναισθήματα στους ανθρώπους ως προβλήματα αναγνώρισης προτύπων και προτάσσονται μοντέλα για την υλοποίηση αυτών των πρακτικών. Τέλος, αναλύονται τα ηθικά διλήμματα στην δόμηση υπολογιστών με συναίσθημα και οι πολλαπλές εφαρμογές στις οποίες θα μπορούσαν να συμβάλουν.

Όπως και σε άλλες γνωσιακές διεργασίες, το συναίσθημα είχε αγνοηθεί αρκετά στις θεωρίες σχετικά με τη μάθηση, ίσως λόγω της δυσκολίας μέτρησής του. Όσον αφορά όλες τις δραστηριότητες που απαιτούν νοητική απόδοση είναι γνωστό ότι το συναίσθημα αποτελεί καθοριστικό παράγοντα. Ο D. O. Hebb εξετάζοντας την σχέση ανάμεσα στην ενεργοποίηση των νευρώνων και το νευρικό σύστημα εισήγαγε την κλασική ανάποδη U καμπύλη που σχετίζει την απόδοση μάθησης με την συναισθηματική ένταση (arousal). Η καμπύλη αυτή ερμηνεύεται με την παραδοχή ότι μία εύκολα αποκτημένη συνήθεια μπορεί να μορφοποιηθεί εύκολα με έντονη διέγερση, ενώ μία δύσκολα αποκτημένη συνήθεια μπορεί να καθιερωθεί μόνο μέσω σχετικά αδύναμης διέγερσης. Η απόδοση στην μάθηση βάζει αυτής της καμπύλης βρίσκεται στο ελάχιστό της όταν κάποιος ξυπνάει ή υφίσταται υπερβολικά μεγάλη συναισθηματική ένταση. Η μέγιστη απόδοση βελτιστοποιείται σε μία ενδιάμεση κατάσταση [20].

Η Picard αναζητά την προσέγγιση της διαδικασίας μάθησης μέσω συναισθηματικών παραμέτρων. Υποστηρίζει ότι αυτές οι ενισχυμένες θεωρίες θα μπορούσαν όχι απλά να βελτιστοποιήσουν την συναισθηματική κατάσταση ενός μαθητή, αλλά να συντονίσουν μία ισορροπία ανάμεσα στην δυσκολία (δυσφορία) και την επίτευξη (ικανοποίηση). Σε μία διαδικασία μάθησης με έναν πομπό και έναν δέκτη μπορούν να μελετηθούν τα οφέλη της μετάδοσης συναισθημάτων, όπως συμβαίνει με τους δασκάλους που εμπνέουν ενδιαφέρον στους μαθητές τους [38]. Πέρα από την ψυχολογικό κίνητρο για μάθηση, που περιλαμβάνει στις περισσότερες των περιπτώσεων εκκίνηση με περιέργεια και ενθουσιασμό και αύξηση σύγχυσης και άγχους με την αυξανόμενη δυσκολία, το μεταιχμιακό κέντρο του εγκεφάλου είναι καθοριστικό σε διάφορες σημαντικές γνωσιακές λειτουργίες που πλαισιώνουν της διαδικασία της μάθησης, όπως αναλύθηκε παραπάνω.

Στην περίπτωση της μάθησης μέσω υπολογιστή, είτε στην περίπτωση μάθησης ενός νέου λογισμικού ή ενός εκπαιδευτικού προγράμματος, η διαδικασία μάθησης δεν θεωρείται ευχάριστη. Η Picard εδώ εισάγει την υπόθεση δόμησης υπολογιστών που αναγνωρίζουν την συναισθηματική έκφραση του χρήστη. Το επιχείρημα εδώ είναι ότι η μάθηση είναι σχεδόν πάντα αμφίδρομη. Όπως λοιπόν ένας δάσκαλος παρατηρεί και μαθαίνει για την

προσοχή και τις προτιμήσεις ενός μαθητή κατά την αλληλεπίδρασή τους, έτσι και ένας συναισθηματικά επαυξημένος υπολογιστής θα μπορούσε να πράττει ανάλογα.

Πέρα από τη διαδικασία της μάθησης η Picard παρουσιάζει πάνω από πενήντα πιθανές εφαρμογές στην ανάκτηση πληροφορίας, τις τηλεπικοινωνίες, την διασκέδαση, τον σχεδιασμό, την υγεία και την ανθρώπινη διάδραση όπου η συναισθηματική υπολογιστική θα μπορούσε να ωφελήσει. Πιο συγκεκριμένα, αναφέρεται σε υπολογιστές που μπορούν να φορεθούν και να αντιληφθούν πληροφορίες συναισθήματος, πλαισίου συνθηκών και φυσιολογίας. Επομένως δημιουργείται η προοπτική συλλογής δεδομένων για πρόοδο στη γνωστική και συναισθηματική θεωρία, όπως και στην βελτίωση της κατανόησης των παραγόντων που συμβάλουν στην υγεία και το ευ ζην.

Ακόμα, πέρα από τους υπολογιστές που αναγνωρίζουν το ανθρώπινο συναίσθημα, η Picard παρουσιάζει την σημασία των υπολογιστικών συστημάτων που μπορούν να εκδηλώνουν συναισθηματική συμπεριφορά. Το συναίσθημα δεν συνιστά συστατικό στοιχείο μόνο για τη δημιουργική συμπεριφορά στους ανθρώπους, αλλά κατέχει καθοριστικό ρόλο και στην λήψη αποφάσεων. Επομένως, καταλήγει στο συμπέρασμα ότι η δόμηση "ευφυών" υπολογιστικών συστημάτων συνεπάγεται δόμηση υπολογιστικών συστημάτων με συναίσθημα.

Τέλος, πέρα από την παρουσίαση των προτερημάτων της ερευνητικής εγκαινίασης της ανάπτυξης εφαρμογών στην συναισθηματική υπολογιστική τίθεται το ηθικό δίλημμα και οι πιθανοί κίνδυνοι από την δόμηση πιθανά "ευφυών" και "συναισθηματικών" υπολογιστικών συστημάτων σε τέτοιο βαθμό που ίσως οδηγηθούν στην απόσχιση από τους δημιουργούς τους [38].

Μετά την εγκαινίαση του επιστημονικού τομέα η Picard ακολούθησε με περαιτέρω δημοσιεύσεις μέσα σε αυτά τα θεματικά πλαίσια. Από την παρουσίαση προκλήσεων που αναδύονται με την εξερεύνηση του τομέα αυτού [39] και την επεξεργασία σημάτων που περιλαμβάνουν συναισθηματική πληροφορία, όπως τα φυσιολογικά σήματα (αγωγιμότητα δέρματος, πίεση αίματος, ηλεκτρομυογραφήματα (electromyography, EMG)) [19] μέχρι την περαιτέρω αναζήτηση στην επικάλυψη της συναισθηματικής υπολογιστικής με την αλληλεπίδραση ανθρώπου και υπολογιστή [40], η Picard συνέχισε να καινοτομεί και να εξετάζει νέες πτυχές του τομέα αυτού. Σημαντική δημοσίευση αποτελεί το μανιφέστο της συναισθηματικής υπολογιστικής που επιχειρεί την καθιέρωση μίας ενοποιημένης θεωρίας μέσω εργαλείων, τεχνολογιών, μοντέλων και περιβαλλόντων μάθησης, αλλά και ορολογίας [46].

Η έκδοση του πρώτου τεύχους του περιοδικού "IEEE Transactions on Affective Computing" πραγματοποιήθηκε μόλις το 2010, ενώ την εισαγωγή του τεύχους κλήθηκε να γράψει η ιδρύτρια του επιστημονικού αυτού κλάδου περιγράφοντας την πορεία και τις αφορμές μέχρι την εισαγωγή του συναισθήματος στη θεωρία των υπολογιστών, αλλά και τη μετέπειτα πορεία υπό την προσωπική της οπτική [41].

2.2.2 Ανάπτυξη και πρακτικές εφαρμογές

Η θεματική βιβλιογραφική ανάλυση του τομέα της συναισθηματικής υπολογιστικής των Ho et al. [52] αποκαλύπτει το ολοένα και αυξανόμενο ακαδημαϊκό ενδιαφέρον για αυτόν τον τομέα. Παρά το γεγονός ότι τις πρώτες δύο δεκαετίες που ακολούθησαν την εισαγωγή της συναισθηματικής υπολογιστικής στο ερευνητικό τοπίο οι σχετικές δημοσιεύσεις ήταν ελάχιστες (181 δημοσιεύσεις στο διάστημα 1995-2005 και 605 το διάστημα 2006-2015), στο διάστημα ανάμεσα στο 2016 και 2020 η δημοσίευση σχετικών ερευνητικών εργασιών αυξήθηκε εκθετικά. Αυτό αποδίδεται κυρίως στην εστίαση του πολιτικού και εμπορικού ενδιαφέροντος στην αξιοποίηση της επιστήμης αυτής σε επιμέρους εφαρμογές [52]. Όσον αφορά τα σύνολα δεδομένων, στα πρώτα χρόνια ανάπτυξης τα δεδομένα κάθε εφαρμογής προερχόταν από μονοτροπικά (unimodal) μέσα καταγραφής, όπως βίντεο προσώπων,

γαλβανική απόκριση δέρματος (**galvanic skin response, GSR**), ηλεκτροεγκεφαλογραφήματα (**EEG**) και ηλεκτροκαρδιογραφήματα (**ECG**). Παρόλα αυτά τα αποτελέσματα των ερευνών που αξιοποιούν αυτού του είδους τα δεδομένα περιορίζονται λόγω των χαρακτηριστικών τους, όπως για παράδειγμα λόγω της απουσίας βιοδεικτών στο βίντεο, της χαμηλής χωρικής ανάλυσης στο ηλεκτροεγκεφαλογράφημα, και της χαμηλής χρονικής ανάλυσης στην γαλβανική απόκριση δέρματος, το ηλεκτροεγκεφαλογράφημα και το ηλεκτροκαρδιογράφημα. Στη συνέχεια, με την εξέλιξη της τεχνολογίας προέκυψε η αξιοποίηση πολυτροπικών (**multimodal**) μέσων στην κατεύθυνση της συναισθηματικής υπολογιστικής. Αυτού του είδους το περιεχόμενο δεν αναφέρεται μόνο στο βίντεο, τον ήχο και το κείμενο, αλλά και στον τομέα των βιολογικών αισθητηριακών καταγραφών του ηλεκτροεγκεφαλογραφήματος, της γαλβανικής απόκρισης δέρματος κτλ. Σε αυτό το πλαίσιο, με την εξέλιξη της τεχνητής νοημοσύνης πολλές εφαρμογές της έχουν αξιοποιηθεί σε ερευνητικές απόπειρες στον τομέα της συναισθηματικής υπολογιστικής. Ένα παράδειγμα αποτελεί η παλινδρόμηση διανυσμάτων υποστήριξης (**support vector regression**) ως ένας πολύ διαδεδομένος τρόπος για την εκτίμηση συναισθήματος. Ανάμεσα σε άλλες εφαρμογές οι κυριότερες είναι η επίταξη της βαθιάς μάθησης στα πλαίσια της συναισθηματικής υπολογιστικής, η ανάλυση και η κατηγοριοποίηση συναισθήματος (**sentiment analysis/ sentiment categorization**). Παραδείγματα τέτοιων εφαρμογών αποτελούν διαδραστικά νευρωνικά δίκτυα για εντοπισμό συναισθήματος σε διαλόγους, πολυτροπική (**multimodal**) ανάλυση συναισθήματος με ιεραρχική συγχώνευση (**hierarchical fusion**) και μοντελοποίηση περιβάλλοντος, όπως και πολυεπίπεδη ανάλυση συναισθήματος με μηχανισμό πολλαπλής προσοχής [18].

Τα πεδία εφαρμογών της συναισθηματικής υπολογιστικής είναι αρκετά και περιλαμβάνουν ανάμεσα σε άλλα τους τομείς της υγείας, της εκπαίδευσης, του μάρκετινγκ και της ασφαλούς οδήγησης. Στον τομέα της υγείας η εξέλιξη του διαδικτύου των ιατρικών πραγμάτων (**Internet of Medical Things, IoMT**) παρέχει μία βάση για αναγνώριση συναισθήματος βάσει φυσιολογικών σημάτων, ενώ υπάρχουν ακόμα σενάρια για τη μελέτη του ρόλου του συναισθήματος στο τομέα της ψυχικής υγείας, της αξιολόγησης φυσικής κατάστασης και της παρακολούθησης ασθενών [42]. Μία σημαντική εφαρμογή στον τομέα της υγείας αποτελεί η βελτίωση της ποιότητας ζωής των ανθρώπων με αυτισμό. Ο αυτισμός χαρακτηρίζεται από νευροαναπτυξιακές συνθήκες οι οποίες οδηγούν σε περιορισμένη κοινωνική διάδραση και δυσκολίες στην επικοινωνία. Έχει παρατηρηθεί, ωστόσο, ότι άτομα που έχουν διαγνωστεί στο φάσμα του αυτισμού διαθέτουν πολύ μεγάλη προσοχή σε λεπτομέρειες, συγκεκριμένα αλλά εμπεριστατωμένα ενδιαφέροντα και ιδιαίτερες ικανότητες (**islets of ability**). Σε αυτή τη βάση ερευνητικές εργασίες περιγράφουν εφαρμογές εντοπισμού και αναγνώρισης συναισθημάτων ως βοηθητικές στην κοινωνική τους λειτουργία, αλλά και στη διαδικασία μάθησης [43].

Στην περίπτωση της ασφαλούς οδήγησης η αναγνώριση συναισθήματος στα πλαίσια της συναισθηματικής υπολογιστικής μπορεί να θεωρηθεί σημαντική για την αποφυγή ατυχημάτων, αφού η αναγνώριση συναισθημάτων επεκτείνεται και στην αναγνώριση του τρόπου οδήγησης. Πιο συγκεκριμένα, αρνητικά συναισθήματα όπως θυμός, φόβος και στεναχώρια επιδρούν αρνητικά στον χρόνο απόκρισης του οδηγού και μπορούν να οδηγήσουν σε θανατηφόρα ατυχήματα.

Στον τομέα της εκπαίδευσης υπάρχει δυναμικό για πολλές εφαρμογές συναισθηματικής υπολογιστικής. Ηλεκτροεγκεφαλογραφήματα, εκφράσεις προσώπου, αλλά και δεδομένα από συσκευές που φοριούνται (**wearable Commercial-Off-The-Shelf, COTS**) παρέχουν πληροφορίες για την συναισθηματική κατάσταση των μαθητών και, κατ'επέκταση, την δυναμική τους μέσα σε ένα περιβάλλον μάθησης [42]. Μία τέτοια εφαρμογή στα πλαίσια μίας "έξυπνης" αίθουσας διδασκαλίας πραγματοποιείται από τους Popescu et al. Οι μετρήσεις πραγματοποιούνται με αισθητήρες, όπως αισθητήρες πίεσης καρτέλας, μέτρησης καρδιακού παλμού, γαλβανικής αγωγιμότητας δέρματος και καταγραφής μέσω καμερών web.

Οι μετρήσεις αυτές και τα συμπεράσματα που προκύπτουν από αυτές θέτουν τις βάσεις για την τροποποίηση εκπαιδευτικών διαδικασιών, έτσι ώστε να αυξηθεί η κινητοποίηση, η προσοχή και η συμμετοχή των μαθητών [44].

Τέλος, ο τομέας της διαφήμισης και του μάρκετινγκ διαθέτει μεγάλο περιθώριο έρευνας όσον αφορά την αναγνώριση συναισθήματος και προσωπικότητας ενός πελάτη. Σε αυτή την κατεύθυνση μπορεί να επιτευχθεί βελτιστοποίηση της αλυσίδας παροχής και ζήτησης, ενώ το κόστος παραγωγής μπορεί να περιοριστεί σημαντικά. Με την συλλογή δεδομένων προτίμησης και απέχθειας (**like/dislike**) για κάποιο προϊόν ή κάποια επωνυμία, αλλά και δεδομένων σημάτων ηλεκτροεγκεφαλογραφήματος και εκφράσεων προσώπου σε συνάρτηση με διαφημίσεις μπορούν να εξαχθούν σημαντικά συμπεράσματα για την στρατηγική μίας εταιρείας και το επικοινωνιακό της πλαίσιο [42].

Κεφάλαιο 3

Συναισθηματικά επαυξημένη τεχνητή νοημοσύνη (**Emotion-augmented AI**)

Θεμέλιο της ενσωμάτωσης του συναισθήματος στον τομέα της επιστήμης των υπολογιστών συνιστά η συναισθηματική υπολογιστική, η οποία εισήγαγε την ανάλυση και τη σύνθεση αντίστοιχων υπολογιστικών μοντέλων. Στις επόμενες παραγράφους παρουσιάζονται οι σημαντικότεροι τρόποι ενσωμάτωσης του "τεχνητού" συναισθήματος σε συστήματα τεχνητής νοημοσύνης με εγγενή τρόπο [16].

3.1 Συναίσθημα σε διεργασίες βελτιστοποίησης

Η συναισθηματική μάθηση στον ανθρώπινο εγκέφαλο αποτελεί μίας μορφής εξαρτημένη μάθηση που συσχετίζεται με εξωτερικά συναισθηματικά ερεθίσματα, όπως επιβραβεύσεις και τιμωρίες που λαμβάνουν χώρα κάτω από διάφορες πραγματικές συνθήκες. Αυτά τα συναισθηματικά ερεθίσματα μπορούν να προκαλέσουν διάφορες εσωτερικές συναισθηματικές καταστάσεις όπως χαρά, λύπη και φόβο, επηρεάζοντας μελλοντικές αποφάσεις. Τα τελευταία χρόνια έχουν εμφανιστεί διάφορα μοντέλα που προσομοιώνουν την διαδικασία συναισθηματικής μάθησης των θηλαστικών (**brain emotional learning, BEL**) και αξιοποιούνται σε προβλήματα ελέγχου και προβλέψεων. Σημαντική συμβολή προς αυτή την κατεύθυνση αποτελεί το έργο των **Morén** και **Balkenius** που εισάγει ένα υπολογιστικό μοντέλο με συναισθηματική μάθηση και λειτουργία εμπνευσμένη από την νευροφυσιολογία του εγκεφάλου των θηλαστικών [4].

Το συναίσθημα έχει φτάσει να εφαρμόζεται σε βαθύτερο επίπεδο στα υπολογιστικά συστήματα με αποτέλεσμα συναισθηματικές έννοιες να έχουν προσαρμοστεί σε δομικό επίπεδο. Με την ερευνητική του εργασία ο **A. Khashman** τροποποιεί ένα πολυεπίπεδο νευρωνικό δίκτυο (**Multi-Layer Perceptron, MLP**) ώστε να εισάγει την ψυχολογική έννοια του άγχους (**anxiety**) και της αυτοπεποίθησης (**confidence**) [24]. Ο αλγόριθμος που εισάγει ονομάζεται συναισθηματική οπισθοδιάδοση (**emotional backpropagation, EmBP**) και περιλαμβάνει μία επιπλέον συναισθηματική πόλωση (**emotional bias**) που εφαρμόζεται στους κρυμμένους νευρώνες και τους νευρώνες εξόδου του δικτύου. Με την εφαρμογή του αλγόριθμου στην αναγνώριση προσώπων το δίκτυο δέχεται ως είσοδο το γενικευμένο μέσο όρο των προτύπων εισόδου που στοχεύει να αποδώσει την πληροφορία κάθε εικόνας σε γενικότερο πλαίσιο. Η επίδραση του άγχους (**anxiety**) συνίσταται στη βελτιστοποίηση της συμπερίληψης του τελευταίου σφάλματος που πραγματοποιήθηκε από το δίκτυο, σε αντίθεση με την αυτοπεποίθηση (**confidence**) που αποδίδει μεγαλύτερη σημασία στις προηγούμενες τροποποιήσεις των παραμέτρων του δικτύου. Για το λόγο αυτό η παράμετρος του άγχους μπορεί να θεωρηθεί ως μοντελοποίηση ενός χαρακτηριστικού εστίασης του δικτύου στο να "μαθαίνει" νέα και ενδιαφέροντα για την εκάστοτε διεργασία δεδομένα.

Σε μετέπειτα ερευνητική δημοσίευση ο A. Khashman ενισχύει τον EmBP αλγόριθμο με την εισαγωγή της εντύπωσης (*impression*). Στόχος με την εισαγωγή αυτού του χαρακτηριστικού είναι το ίδιο να δράσει ως ένα οπτικό σήμα για το αντικείμενο προς αναγνώριση που να ενισχύει ή να αποδυναμώνει την παράμετρο του άγχους ανάλογα την θετική ή αρνητική αντίδραση. Η παράμετρος της εντύπωσης (*impression*) προκύπτει από τη μέση τιμή των γκρι επιπέδων της εικόνας, με τις πιο σκούρες εικόνες να προκαλούν χειρότερες εντυπώσεις. [25]

Σε ένα διαφορετικό μοντέλο και μία νέα ερευνητική εργασία ο A. Khashman εισάγει ένα νέο είδος νευρωνικού δικτύου, το DuoNN. Το DuoNN (Duo Neural Network) αποτελείται από τους DuoNeurons που απαρτίζονται από έναν γνωσιακό (*cognitive*) και έναν συναισθηματικό νευρώνα (*emotional*). Ο πρώτος εστιάζει στα τοπικά χαρακτηριστικά, ενώ ο δεύτερος επεξεργάζεται γενικότερα μοτίβα. Το μοντέλο αυτό εμπνέεται από τον οπτικό φλοιό (*visual cortex*) και την κατάτμηση της ροής πληροφορίας σε ραχιαίο (*dorsal*)/γνωσιακό (*cognitive*) μονοπάτι και σε κοιλιακό (*ventral*)/συναισθηματικό (*affective*) μονοπάτι [26].

Σε συνέχεια ανάπτυξης ενός αλγορίθμου βελτιστοποίησης εμπνευσμένου από τη θεωρία του χάους που βρέθηκε να συγκλίνει σε τοπικά ελάχιστα κατά τη διάρκεια διεργασιών βελτιστοποίησης, οι Yang et al. απέδειξαν ότι η εισαγωγή τεχνητού συναισθήματος μπορεί να οδηγήσει σε αποφυγή τοπικών ελαχίστων [60]. Το μοντέλο που παρουσιάζεται απεικονίζει τον ρόλο του συναισθήματος κατά την εξερεύνηση-εκμετάλλευση (*exploration-exploitation*) και η τεχνική υλοποίησή του περιλαμβάνει αισθήσεις, αισθήματα και συναισθήματα, όπως και σε ένα ορμονικό σύστημα που ανατροφοδοτεί την ενεργοποίηση αισθημάτων. Εδώ τα συναισθήματα μοντελοποιούνται περισσότερο αφαιρετικά και περιλαμβάνουν το άγχος (*anxiety*) που ενεργοποιείται με φθίνουσα επιβράβευση, την αυτοπεποίθηση (*confidence*) που αποτελεί αρνητική συνάρτηση του άγχους, το φόβο που αυξάνεται ανάλογα του άγχους και εμπλέκεται σε στρατηγικό έλεγχο και την θέρμη (*warmth*) που λειτουργεί ως σηματοδότης του τέλους του αλγορίθμου βάσει του φόβου και του αριθμού των επαναλήψεων. Με την μοντελοποίηση αυτών των συναισθημάτων προκύπτει ένα κυρίαρχο συναίσθημα το οποίο εφαρμόζεται ως το επικρατέστερο επηρεάζοντας τις αλγοριθμικές παραμέτρους και επιφέροντας στρατηγικές/συμπεριφορικές αλλαγές. Οι Yang et al. ανακάλυψαν ότι με την χρήση τεχνητού συναισθήματος η ταχύτητα σύγκλισης αυξήθηκε και ξεπέρασε τις συγκρινόμενες μεθόδους.

Αξίζει να σημειωθεί ένα ακόμα επαυξημένο MLP μοντέλο που διαθέτει νευρορρυθμιστικούς παράγοντες που προσομοιάζουν ορμονικούς αδένες. Το δίκτυο αυτό ονομάζεται EMANN (EMotional Artificial Neural Networks) [45]. Οι αδένες αυτοί διαχειρίζονται από το νευρωνικό δίκτυο και είναι στην ουσία νευρώνες μέσα σε αυτό. Επομένως οι νευρώνες είναι παραλήπτες και αποστολείς ορμονών. Αυτές οι ορμόνες επηρεάζουν το νευρωνικό δίκτυο με την τροποποίηση του κατωφλιού ενεργοποίησης (*activation threshold*), τον υπολογισμό του δυναμικού εισόδου του νευρώνα, αλλά και της εξόδου του. Η απόδοση του μοντέλου αυτού φαίνεται να προσπερνά σε μεγάλο βαθμό την απόδοση του απλού MLP δικτύου, αφού μετά το πέρας της περιόδου αύξησης της απόδοσης και των δύο δικτύων το νέο μοντέλο συνέχισε να βελτιώνει την ορθότητα (*accuracy*) του σε αντίθεση με το μοντέλο αναφοράς. Συμπεραίνεται λοιπόν στην ερευνητική αυτή η εργασία ότι οι ορμονικές συνέπειες των συναισθημάτων είναι ένας πολλά υποσχόμενος σχεδιασμός στην τεχνητή νοημοσύνη.

3.2 Ανατομικά μοντέλα στην τεχνητή νοημοσύνη

Πέρα από τα μοντέλα που προσεγγίζουν συγκεκριμένες διαστάσεις του συναισθήματος όσον αφορά την αλλαγή στη συμπεριφορά ενός ανθρώπου, έχει πραγματοποιηθεί έρευνα πάνω σε ολιστικά μοντέλα που μοντελοποιούν την ανατομία του εγκεφάλου. Ένα παράδειγμα τέτοιας εγκεφαλικής δομής προς μοντελοποίηση είναι το μεταιχμιακό σύστημα (limbic system) [16]. Οι Morén και Balkenius κατασκεύασαν υπολογιστικά μοντέλα με δομικά στοιχεία που αναπαριστούν τον ιππόκαμπο (hippocampus), τον κογχομετωπιαίο φλοιό (orbitofrontal cortex, OFC), την αμυγδαλή (amygdala), τον θάλαμο (thalamus), και τον κινητικό φλοιό (motor cortex) με σκοπό την αναπαραγωγή συναισθημάτων βάσει των εξαρτήσεων ανάμεσα σε αυτές τις εγκεφαλικές δομές [4]. Η αμυγδαλή μαθαίνει συναισθηματικούς συσχετισμούς, ενώ ο κογχομετωπιαίος φλοιός λειτουργεί ως αναστολέας συνθηκών. Οι συνθήκες αυτές παρέχονται ως μία έγχυση στον κογχομετωπιαίο φλοιό από τον ιππόκαμπο, όπου και παράγεται μέσω της αντιστοίχισης ερεθισμάτων σε τοποθεσίες. Αυτές οι αντιστοιχίσεις δημιουργούν μία προσδοκία για ερέθισμα, το οποίο αν δεν εμφανιστεί προκύπτει μεταβολή στις συνθήκες.

Η έρευνα των Morén και Belkenius έθεσε ένα πρωτοπόρο υπολογιστικό πλαίσιο που αξιοποιεί την αμυγδαλή και τα σχετιζόμενα συστήματα του εγκεφάλου. Οι Lucas et al. [7] δημιούργησαν έναν αλγόριθμο εμπνευσμένο από το μεταιχμιακό σύστημα που ονομάστηκε BELBIC και χρησιμοποιήθηκε ως ελεγκτής σε μηχανικές διεργασίες. Σε σύγκριση με έναν αναλογικό αναπόσπαστο παράγωγο ελεγκτή (proportional-integral-derivative, PID) και βελτιστοποιημένους γραμμικούς ελεγκτές, το BELBIC παρουσιάζει καλύτερες δυνατότητες βελτιστοποίησης και προσαρμοστικότητας με την προσαρμογή σε παραμετρικές αλλαγές και αλλοιώσεις. Εδώ ενσωματώνεται η ιδέα του σύντομου και μακρύτερου μονοπατιού ερεθίσματος προς την αμυγδαλή, όπου το πιο σύντομο μονοπάτι γεφυρώνει τον αισθητικό θάλαμο με την αμυγδαλή ενώ το μακρύτερο μονοπάτι περνάει από το μετωπιαίο φλοιό.

Οι Babaie et al. ανέπτυξαν το BEL (Brain Emotional Learning, BEL) μοντέλο [50]. Το μοντέλο αυτό βάσει συγκρίσεων της εξόδου της αμυγδαλής με τις επιβραβεύσεις από τα δεδομένα εισόδου ορίζει το κογχομετωπιαίο φλοιό ως ρυθμιστή. Πιο συγκεκριμένα, στην περίπτωση που υπάρχει επιβράβευση ο κογχομετωπιαίος φλοιός ευθυγραμμίζει την έξοδο της αμυγδαλής ή απορρίπτει την έξοδο της στην περίπτωση που δεν υπάρχει επιβράβευση. Το BEL μοντέλο ξεπέρασε σε απόδοση το κλασικό MLP και το ANFIS μοντέλο (Adaptive Network based Fuzzy Inference System) [23].

Οι Lotfi και Akbarzadeh-T ανέπτυξαν ένα νέο μοντέλο προσαρμόζοντας τα συμπεράσματα από την έρευνα των Morén και Balkenius [12]. Το μοντέλο αυτό ονομάζεται BELPR (Brain Emotional Learning-based Pattern Recogniser) και κατάφερε να ξεπεράσει το MLP σε ταξινομήσεις πολλαπλών εισόδων και εξόδων και σε προβλέψεις χαοτικών χρονοσειρών. Σε ένα διαφορετικό μοντέλο, οι Lotfi και Akbarzadeh-T εισάγουν μία διαδικασία "λήθης" για την αμυγδαλή. Το μοντέλο αυτό ονομάζεται ABDEL (Adaptive Decayed Brain Emotional Learning) και χρησιμοποιεί έναν φθίνων ρυθμό που καθορίζει την ταχύτητα με την οποία τα βάρη της αμυγδαλής περιορίζουν το ποσοστό των προηγούμενων τιμών των βαρών που περιέχονται στην τρέχουσα τιμή τους. Η λογική αυτή εμπνέεται από την ανάγκη για ένα σύστημα που προσαρμόζεται σε ένα μεταβαλλόμενο περιβάλλον. Το ABDEL ξεπερνά σε απόδοση τα ANFIS και MLP και εμφανίζει προσαρμοστικότητα σε online προβλέψεις. Μία διαφοροποίηση που έχει καταφέρει η ίδια ερευνητική ομάδα είναι η εισαγωγή ασαφούς λογικής στο μοντέλο αυτό, όπου οι μεταβλητές εισόδου και εξόδου της αμυγδαλής και του κογχομετωπιαίου λοβού μετατρέπονται σε ασαφείς και η αξιολόγηση πραγματοποιείται με ασαφείς τελεστές [13]. Η προσθήκη αυτή αποδεικνύει ότι η εισαγωγή ασαφούς λογικής προκαλεί ταχύτερη σύγκλιση σε σύγκριση με τα MLP και ADBEL μοντέλα και ότι ένα συναισθηματικά επαυξημένο μοντέλο μπορεί να λειτουργήσει ως ένας ασαφής αλγόριθμος.

Οι Lotfi και Akbarzadeh-T πρότειναν επίσης το μοντέλο LiAENN (*Limbic-Based Artificial Emotion Neural Network*) που, όπως το όνομά του προδίδει, βασίζεται στο μεταιχμιακό σύστημα (*limbic system*) [8]. Σε αυτή τους την προσέγγιση συνδυάζονται οι συναισθηματικοί παράμετροι του άγχους (*anxiety*) και της αυτοπεποίθησης (*confidence*) με τα συντομότερα και μακρύτερα μονοπάτια ροής της πληροφορίας, τις διεργασίες "λήθης" της αμυγδαλής, και τον περιορισμό συναισθήματος μέσω της αλληλεπίδρασης ανάμεσα στον κορχομετωπιαίο φλοιό και την αμυγδαλή. Αυτός ο αλγόριθμος είναι ένας συγκερασμός μεταξύ του EmBP, όπου οι συναισθηματικοί παράμετροι αξιοποιούνται στη μάθηση, και των εγκεφαλικών δομών που εμπνέονται από όλα τα μοντέλα που αναφέρθηκαν παραπάνω. Το μοντέλο αυτό καταφέρνει να ξεπεράσει σε απόδοση το μοντέλο αμυγδαλής και κορχομετωπιαίου φλοιού των Morén και Balkenius, όπως και τα BELBIC, EmBP, DuoNN και BELPR. Συμπεραίνεται επομένως ότι ο συνδυασμός πολλών βελτιστοποιήσεων που οδηγούν σε συναισθηματικά επαυξημένα δίκτυα τεχνητής νοημοσύνης παράγει και καλύτερα αποτελέσματα.

3.3 Συναίσθημα στην υπολογιστική ενισχυτική μάθηση

Η εισαγωγή συναισθήματος στην υπολογιστική ενισχυτική μάθηση (*Reinforcement Learning, RL*) έχει αρχίσει να εμφανίζεται και να παράγει πιο εκλεπτισμένα μοντέλα. Υπάρχουν διαφορετικού είδους προσεγγίσεις στον συνδυασμό συναισθήματος και υπολογιστικής ενισχυτικής μάθησης, όπου ο πράκτορας (*agent*) πρέπει να μάθει την βέλτιστη πολιτική αντιστοίχισης κατάστασης-δράσης και της μέγιστης επιβράβευσης. Μία προσέγγιση προκύπτει από την θεωρία αξιολόγησης συναισθήματος (*appraisal theory*), όπου αξιολογήσεις πραγματοποιούνται βάσει τη κατάστασης και οι επιβραβεύσεις προσαρμόζονται αντίστοιχα. Μία σημαντική προσαρμογή για αυτές τις μεθόδους είναι η σύλληψη εγγενούς κινητοποίησης (σχετιζόμενη με εσωτερική ή συναισθηματική επιβράβευση) και εξωγενούς κινητοποίησης (η εξωτερική επιβράβευση είναι κοινή στην τυπική ενισχυτική μάθηση) στην επιβράβευση της ενισχυτικής μάθησης. Ένα τέτοιο παράδειγμα από τους Marinier και Laird δείχνει ότι η ενσωμάτωση συναισθήματος που προκύπτει από αξιολόγηση οδηγεί στην ταχύτερη μάθηση σε σύγκριση με έναν κλασικό πράκτορα (*agent*) ενισχυτικής μάθησης [17]. Αυτό αποδίδεται στο ότι το συναίσθημα διαθέτει συνεχή επιβράβευση σε αντίθεση με την βασική ενισχυτική μάθηση όπου η επιβράβευση λαμβάνεται μόνο στο τέλος κάθε επανάληψης. Επιπρόσθετα, με την προσθήκη της παραμέτρου της διάθεσης (*mood*), η οποία διατηρεί το συναίσθημα για κάθε εκτέλεση, η διαδικασία της μάθησης επιταχύνεται περισσότερο. Αυτό οφείλεται στο γεγονός ότι κάποιες απαραίτητες ενέργειες δεν προκαλούν κανένα συναισθηματικό αποτέλεσμα. Χωρίς την επίδραση της παραμέτρου της διάθεσης αυτές οι καταστάσεις μπορεί να υποτιμηθούν, αλλά με την διατήρηση της συναισθηματικής επιβράβευσης η αξία κάθε κατάστασης προσεγγίζεται με μεγαλύτερη επιτυχία. Άλλες παρόμοιες θεωρίες βάσει αξιολόγησης συναισθημάτων έχουν αποδειχθεί ότι προκαλούν καλύτερα αποτελέσματα σε σχέση με την απλή ενισχυτική μάθηση [36], [37].

Εναλλακτικά, κάποια από τα μοντέλα της συναισθηματικής ενισχυτικής μάθησης περιλαμβάνουν την έννοια της ομοιότητας και του κινήτρου, δηλαδή της έλξης προς ένα σημείο ισορροπίας του πράκτορα (*agent*). Σε αυτά τα μοντέλα η ομοιότητα περιγράφει μία κατάσταση ισορροπίας για τον πράκτορα ως προς τις ομοιοστατικές ή βασικές μεταβλητές, όπως η πείνα και η δίψα. Ο πράκτορας βιώνει κίνητρα που προκαλούν δράσεις και μεταβάλλουν την κατάστασή του σε ένα ομοιοστατικό χώρο προς το σημείο ισορροπίας (*equilibrium*). Αυτές οι αλλαγές στη συνέχεια μπορούν να μεταφραστούν σε επιβραβεύσεις.

Ακόμα, το συναίσθημα και άλλες αρχές εμπνευσμένες από την νευροβιολογία έχουν ενσωματωθεί στην επιλογή δράσης. Για παράδειγμα, η δραστικότητα (*valence*) ενός συναι-

σθήματος έχει αξιοποιηθεί στην επιρροή της εξερεύνησης-εκμετάλλευσης (**exploration-exploitation**), όπου η αρνητική ή η θετική δραστηριότητα προωθεί αντίστοιχα περισσότερο ή λιγότερο την αξιολόγηση στην επιλογή δράσης. Σε ένα διαφορετικό παράδειγμα η παράμετρος της αγανάκτησης (**frustration**) χρησιμοποιείται για να ορίσει ποια στρατηγική ή ώθηση θα πρέπει να ληφθεί υπόψη για την επιλογή μίας ενέργειας σε ένα τροφοσυλλεκτικό ρομποτικό περιβάλλον (π.χ. εναλλαγή ανάμεσα στην οδομετρική πλοήγηση και την οπτική πλοήγηση ή στην παραχώρηση προτεραιότητας στο κίνητρο της δίψας ή της πείνας) [5]. Με ανάλογο τρόπο η εισαγωγή της παραμέτρου της αγανάκτησης ως μέρος του μοντέλου της αμυγδαλής στο **Q-learning** αποδείχθηκε να βελτιώνει την απόδοση της πλοήγησης ενός πράκτορα, όπου το συναίσθημα τροποποιεί αντίστοιχα τα βάρη διαφορετικών συναρτήσεων αξίας που αξιολογούν και οδηγούν στην επιλογή ενέργειας [49].

3.4 Νόηση και αφαιρετικότητα στη μάθηση

Σε ένα περιβάλλον προσομοίωσης ρομποτικών εφαρμογών οι **Parisi** και **Petrosino** [6] ανακάλυψαν ότι η προσθήκη ενός "συναισθηματικού κυκλώματος" σε έναν πράκτορα ελεγχόμενο από νευρωνικό δίκτυο επιτρέπει την υλοποίηση κινητήριας λήψης αποφάσεων. Η αρχιτεκτονική των νευρωνικών δικτύων δημιουργήθηκε με έναν γενετικό αλγόριθμο και βρέθηκε ότι ρομποτικά συστήματα που περιλαμβάνουν συναίσθημα επιτυγχάνουν μεγαλύτερη προσαρμογή όταν αξιολογούνται σε έναν αριθμό περιβαλλόντων και με διαφορετικά κίνητρα. Για παράδειγμα, κάποια συστήματα θα πρέπει να αναζητήσουν φαγητό και νερό και το επιλεγμένο περιβάλλον θα πρέπει να παρέχει διάφορες ποσότητες καθενός από αυτά τα αγαθά. Σε μία περίπτωση η ποσότητα του φαγητού επιλέγεται να είναι πέντε φορές μεγαλύτερη από αυτή του νερού. Αυτό οδηγεί τα ρομποτικά συστήματα να αποδίδουν μεγαλύτερη προτεραιότητα στο νερό, αναζητώντας φαγητό μόνο όταν βρίσκονται κοντά του ή όταν τα επίπεδα πείνας έχουν αυξηθεί πολύ. Ο **Velásquez** εισάγει τον έλεγχο βάσει συναισθήματος (**emotion-based control**) [54], όπου η επιλογή δράσης πραγματοποιείται βάσει συναισθηματικών συσχετίσεων στη μάθηση και οδηγούν μετέπειτα αποφάσεις μέσω σημάτων πώλωσης (**bias**). Με τη χρήση του μοντέλου Κάθεξις (**Cathexis**) που αποτελείται από ένα σύνολο συστημάτων που περιλαμβάνουν μία γεννήτρια συναισθημάτων (**emotion generator**), συμπεριφορικά συστήματα (**behaviour systems**) και κινητικά συστήματα (**motor systems**) ο **Velásquez** συνδυάζει την ικανότητα μάθησης με κάποιους βασικούς κανόνες για απόσπαση συναισθήματος. Αυτός ο συγκεκρισμός υλοποιείται σε έναν ρομποτικό πράκτορα που ονομάζεται "**Yuppy**" και επιδεικνύει συναισθηματική μάθηση (**emotional conditioning**). Ο **Yuppy** μαθαίνει με το πέρασ του χρόνου συναισθηματικούς συσχετισμούς, όπως τον συσχετισμό του να είναι πειθαρχημένος μέσω του φόβου ή τον συσχετισμό μέσω της δειλιάς ανάμεσα στον φόβο και έναν ήχο για τον οποίο προηγουμένως δεν διαθέτει κανέναν γνωσιακό συσχετισμό. Σε διαφορετική περίπτωση ο **Yuppy** θα βίωνε χαρά βρίσκοντας ένα ροζ κόκαλο βάσει βασικών κανόνων. Παρόλα αυτά, εαν ένα άτομο κουβαλούσε ένα κόκαλο και πειθαρχούσε το ρομπότ, ο **Yuppy** θα εξελισσόταν επιφυλακτικός προς τους ανθρώπους σε μετέπειτα δοκιμές.

Σε μία διαφορετική γνωσιακή αρχιτεκτονική ο κατανεμημένος πράκτορας με νοημοσύνη μάθησης (**LIDA**) συνδυάζει προσοχή, επιλογή δράσης και κίνητρο με συναισθήματα [33]. Τα συναισθήματα παράγονται ως αποτέλεσμα των αξιολογήσεων γεγονότων (παράγοντας συναισθηματικής συσχέτισης), τα οποία με τη σειρά τους αποτελούν κίνητρο για διάφορες ενέργειες. Στο **LIDA** το συναίσθημα βελτιώνει επίσης τη μάθηση μέσω έντασης (**arousal**) συναισθήματος, αλλά αν η ένταση μεγαλώσει αρκετά προκαλεί παρεμβολές. Αυτές οι πρωτοπόμενες προσεγγίσεις της αφαιρετικής ενισχυτικής μάθησης ανοίγουν το δρόμο για την παραγωγή νέων συμπεριφορών από αρχικούς βασικούς κανόνες που μπορούν να εξελιχθούν αφαιρετικά. Με τον τρόπο αυτό η κατανόηση της θετικής και της αρνητικής

συμπεριφοράς μπορεί να εξελιχθεί σε πολύ περισσότερο εκλεπτυσμένα μοντέλα. Αυτού του είδους η προσέγγιση είναι ευθυγραμμισμένη τόσο με τη θεωρία αξιολόγησης συναισθήματος (*appraisal theory*), όσο και με την αντίληψη ότι ο ρόλος των συναισθημάτων είναι να εξυπηρετήσουν τις βασικές ανάγκες ενός οργανισμού παράγοντας νέους στόχους που ευθυγραμμίζονται με αυτές τις ανάγκες.

Κεφάλαιο 4

Υβριδικό δίκτυο συναισθηματικής μάθησης

Το νευρωνικό δίκτυο που επιλέχθηκε να υλοποιηθεί συνδυάζει την συναισθηματική μάθηση, που πραγματοποιεί ο αλγόριθμος του EmBP, με χάρτες αυτο-οργάνωσης για την προσομοίωση της αντίληψης της οπτικής πληροφορίας από τα συναισθηματικά (κοιλιακά/ventral) δίκτυα στον εγκέφαλο.

4.1 Χάρτες αυτο-οργάνωσης (Self-Organizing Maps, SOM)

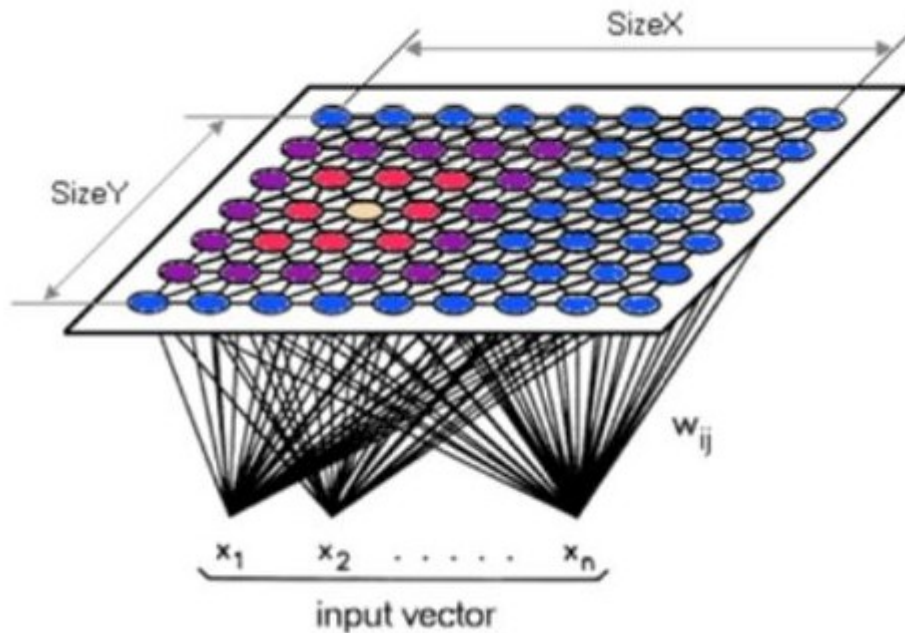
Σημαντικό δομικό στοιχείο στο επιλεγμένο υβριδικό νευρωνικό δίκτυο αποτελούν οι χάρτες αυτο-οργάνωσης ή Self-Organizing Maps (SOM). Για το λόγο αυτό παρουσιάζεται εδώ μία σύντομη περιγραφή της δομής αυτού του δικτύου, της συμπεριφοράς του, και των συνηθισμένων εφαρμογών του.

4.1.1 Γενικά

Οι χάρτες αυτο-οργάνωσης (SOM) δημιουργήθηκαν στην αρχή της δεκαετίας του 1980 από τον Φινλανδό καθηγητή και ερευνητή **Teuvo Kohonen** [29]. Οι χάρτες αυτοί συνιστούν ένα μοντέλο μη επιβλεπόμενης μηχανικής μάθησης και προορίζονται για εφαρμογές όπου η διατήρηση της τοπολογίας ανάμεσα στους χώρους εισόδου και εξόδου του μοντέλου έχει ιδιαίτερη σημασία. Το πλέον σημαντικό χαρακτηριστικό αυτού του αλγορίθμου είναι ότι πρότυπα εισόδου που βρίσκονται κοντά σε χώρο υψηλών διαστάσεων συνεχίζουν να βρίσκονται κοντά στον δισδιάτατο χώρο που αναπαρίσταται από τον χάρτη. Οι χάρτες αυτο-οργάνωσης επομένως συνιστούν μία μέθοδο μείωσης διαστάσεων, αφού πραγματοποιείται η αντιστοίχιση ανάμεσα σε ένα χώρο εισόδου μεγάλων διαστάσεων και έναν χώρο συνήθως χαμηλών διαστάσεων διακριτής αναπαράστασης, όπου η δομή των δεδομένων εισόδου διατηρείται [14].

4.1.2 Δομή

Οι χάρτες αυτο-οργάνωσης διαθέτουν δύο επίπεδα, με το πρώτο να αποτελεί το επίπεδο εισόδου και το δεύτερο το επίπεδο εξόδου, ή διαφορετικά διατυπωμένο, τον χάρτη απεικόνισης χαρακτηριστικών (feature map). Οι χάρτες αυτο-οργάνωσης δεν διαθέτουν κάποια συνάρτηση ενεργοποίησης στους νευρώνες που αποτελούν τον χάρτη, ενώ οι συνδέσεις ανάμεσα στο επίπεδο εισόδου και τον χάρτη διαθέτουν βάρη. Συνεπώς, κάθε νευρώνας στο SOM ανατίθεται με ένα διάνυσμα βαρών ίδιας διάστασης με αυτή της εισόδου του δικτύου .



Εικόνα 4.1: Τοπολογία SOM

4.1.3 Μάθηση

Η μάθηση στους χάρτες αυτο-οργάνωσης δεν βασίζεται στη οπισθοδιάδοση (**back-propagation**) για την ανανέωση των βαρών. Αντίθετα, εδώ αξιοποιείται η ανταγωνιστική μάθηση. Οι τρεις βασικές συνιστώσες της ανταγωνιστικής μάθησης είναι οι εξής [28]:

- **Ανταγωνισμός:** Όπως αναφέρθηκε και παραπάνω, σε κάθε νευρώνα του χάρτη χαρακτηριστικών (**feature map**) ανατίθεται ένα διάνυσμα βαρών μήκους ίσου με τη διάσταση εισόδου. Στο στάδιο αυτό υπολογίζεται η απόσταση ανάμεσα στα δεδομένα εισόδου και τον κάθε νευρώνα του χάρτη χαρακτηριστικών. Ο νευρώνας με την μικρότερη απόσταση από τα δεδομένα εισόδου ανακηρύσσεται νικητής νευρώνας (**Best Matching Unit, BMU**). Η μετρική απόστασης που χρησιμοποιείται πιο συχνά είναι η Ευκλείδεια απόσταση. Η εξίσωση που εκφράζει την ανεύρεση του νικητή νευρώνα i είναι η εξής:

$$i(x^{(n)}) = \underset{j}{\operatorname{argmin}} \|x^{(n)} - w_j\|$$

Όπου $x^{(n)}$ το πρότυπο εισόδου και w_j το διάνυσμα βαρών του εκάστοτε νευρώνα.

- **Συνεργασία:** Μετά την ανεύρεση του νικητή νευρώνα το διάνυσμα βαρών του ανανεώνεται. Ωστόσο, ανανεώνονται σε κάποιο βαθμό και οι γειτονικοί του νευρώνες. Η επιλογή των γειτονικών νευρώνων πραγματοποιείται με τη βοήθεια μίας συνάρτησης πυρήνα (**neighborhood kernel function**), η οποία εξαρτάται από δύο παράγοντες. Οι παράγοντες αυτοί είναι ο χρόνος και η απόσταση από τον νικητή νευρώνα. Στην αρχή της διαδικασίας μάθησης η ακτίνα της γειτονιάς είναι αυθαίρετα μεγάλη αλλά μειώνεται στην πορεία. Με τον τρόπο αυτό εξασφαλίζεται ότι στην αρχή της μάθησης λαμβάνει χώρα μία καθολική ταξινόμηση, ενώ προς το τέλος της διαδικασίας μάθησης, που η ακτίνα μειώνεται σημαντικά, πραγματοποιούνται τοπικές και πιο συγκεκριμένες διορθώσεις στα διανύσματα βαρών.

- Προσαρμογή: Μετά τον καθορισμό του νικητή νευρώνα και των γειτονικών του νευρώνων, τα αντίστοιχα βάρη ανανεώνονται. Η ανανέωση αυτή ωστόσο δεν είναι του ίδιου μεγέθους, αφού εξαρτάται και πάλι από την απόσταση του κάθε νευρώνα από τον νικητή. Πιο συγκεκριμένα, η ανανέωση πραγματοποιείται βάσει της ακόλουθης εξίσωσης:

$$w_k = w_k + \eta(t)h_{ik}(t)(x^{(n)} - w_k)$$

Όπου w_k το διάνυσμα βαρών του νευρώνα k , $\eta(t)$ η φθίνουσα συνάρτηση ρυθμού μάθησης, $h_{ik}(t)$ η συνάρτηση γειτνίασης ανάμεσα στον εξεταζόμενο νευρώνα και τον νικητή, $x^{(n)}$ το πρότυπο εισόδου και w_k το διάνυσμα βαρών του εξεταζόμενου νευρώνα.

4.1.4 Σύγκλιση

Η μαθηματική ανάλυση του αλγορίθμου των χαρτών αυτο-οργάνωσης αποδείχθηκε αρκετά δύσκολη. Η απόδειξη της σύγκλισης της διαδικασίας μάθησης των SOM για τους μονοδιάστατους χάρτες πραγματοποιήθηκε το 1987 από τους Cottrell και Fort, αλλά για μεγαλύτερης διάστασης SOM's οι ιδιότητες σύγκλισης έχουν μελετηθεί σε γενικότερο πλαίσιο και έχουν προταθεί κριτήρια όπως το **fConv**, που αποτελεί μία απλή και κομψή μέθοδο για το βαθμό που η πυκνότητα συχνότητας των δεδομένων εισόδου έχουν μοντελοποιηθεί από τον εκάστοτε χάρτη αυτο-οργάνωσης [35].

4.1.5 Εφαρμογές

Οι χάρτες αυτο-οργάνωσης μπορούν να θεωρηθούν από διαφορετικές οπτικές γωνίες λόγω της πολυχρηστικότητάς τους. Μερικές εφαρμογές τους παρατίθενται στην συνέχεια [21]:

1. Πιθανά η πιο τυπική αντίληψη του SOM είναι ως ένα τεχνητό νευρωνικό δίκτυο που προσμοιώνει δομές του εγκεφάλου, ιδιαίτερα δεδομένης της νευρολογικής παραδοχής ότι ο φλοιός του εγκεφάλου αποτελείται από δομές διατεταγμένων "χαρτών". Υπάρχει αρκετή νευροφυσιολογική απόδειξη που υποστηρίζει την ιδέα ότι οι χάρτες αυτο-οργάνωσης συλλαμβάνουν μερικές από τις πιο βασικές αρχές του εγκεφάλου. Μερικά μοντέλα αυτο-οργάνωσης είχαν εμφανιστεί πριν τα SOM, αλλά τα τελευταία φαίνεται να μοντελοποιούν με μεγαλύτερη επιτυχία της εγκεφαλικές δομές.
2. Οι χάρτες αυτο-οργάνωσης μπορούν να θεωρηθούν και ως ένα μοντέλο μη-επιβλεπόμενης μάθησης, αλλά και ως ένα ευπροσάρμοστο γνωσιακής αναπαράστασης δίκτυο. Οι παραδοσιακοί γνωσιακοί φορμαλισμοί αναπαράστασης γνώσης, όπως τα σημασιολογικά δίκτυα και η κατηγορηματική λογική, είναι στατικοί και οι αναφορικές σχέσεις των στοιχείων είναι καθορισμένες από ανθρώπινο παράγοντα. Ακόμα, αυτοί οι φορμαλισμοί βασίζονται στην υπόθεση ότι η σχέση ανάμεσα στη φυσική γλώσσα και τον πραγματικό κόσμο είναι καθορισμένη ένα προς ένα. Πιο συγκεκριμένα, πραγματοποιείται η εικασία ότι ο πραγματικός κόσμος αποτελείται από αντικείμενα και σχέσεις ανάμεσα σε αυτά, ενώ αυτές οι οντότητες έχουν άμεση αντιστοίχιση με τα στοιχεία της γλώσσας. Σχετικά με αυτή την γνωσιακή αναπαράσταση και την μάθηση αυτής, οι γνωσιακές και φιλοσοφικές διαστάσεις είναι πολύ σχετικές.
3. Οι χάρτες αυτο-οργάνωσης μπορούν να αξιοποιηθούν επίσης στην στατιστική ανάλυση και οπτικοποίηση. Πιο συγκεκριμένα, χρησιμοποιούνται συχνά ως ένα στατιστικό εργαλείο για πολυμεταβλητή ανάλυση. Συνιστά ταυτόχρονα μία μέθοδο προβολής από χώρους υψηλών διαστάσεων σε χώρους χαμηλών διαστάσεων, αλλά και μία

μέθοδος συσταδοποίησης που αντιστοιχεί παρόμοια δεδομένα εισόδου σε γειτονικούς νευρώνες.

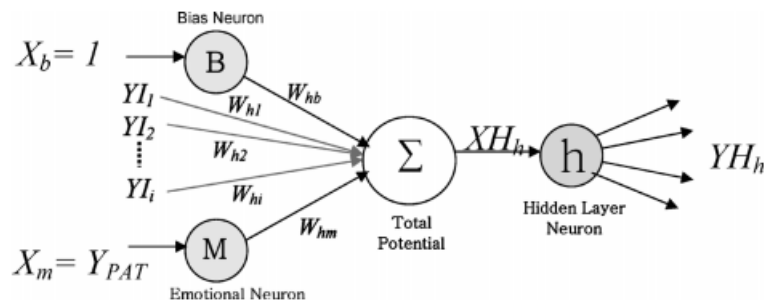
4. Οι δομές **SOM** αποτελούν ένα σημαντικό εργαλείο για την ανάπτυξη πολύπλοκων εφαρμογών. Στον χώρο της εξόρυξης γνώσης από δεδομένα και της οπτικοποίησης της για σύνθετα σύνολα δεδομένων, οι χάρτες αυτο-οργάνωσης μπορούν να εφαρμοστούν σε αναγνώριση ομιλίας, οικονομική ανάλυση, έλεγχο διεργασιών και διαγνωστικές διεργασίες στη βιομηχανία και την φαρμακευτική.

4.2 Συναισθηματικά επαυξημένο **SOM (EmSOM)**

Το επιλεγμένο υβριδικό νευρωνικό δίκτυο που υλοποιείται στην παρούσα εργασία παρουσιάζεται από τους Dai et al., περιλαμβάνει χάρτες αυτο-οργάνωσης και ένα προσθιοδρομικό δίκτυο (*feed-forward neural network*) και ονομάζεται **EmSOM (Emotional SOM)** [9]. Η δομή αυτή εμπνέεται από προγενέστερη έρευνα πάνω σε συναισθηματικά επαυξημένα νευρωνικά δίκτυα που επιχειρούν την μοντελοποίηση συμπεριφορικών αλλά και ανατομικών στοιχείων. Το συγκεκριμένο δίκτυο βασίζεται στην εισαγωγή της συναισθηματικής παραμέτρου στην διαδικασία μάθησης, αλλά και στην ενίσχυση των δεδομένων εισόδου ενός προσθιοδρομικού δικτύου [27], [24].

4.2.1 Θεωρητικό υπόβαθρο

Η συναισθηματικά επαυξημένη οπισθοδιάδοση (*emotional backpropagation, EmBP*) αποτέλεσε βάση για τη εισαγωγή του **EmSOM** δικτύου. Η έμπνευση πίσω από το **EmBP** προέκυψε από την παραδοχή ότι συγκεκριμένα ανθρώπινα συναισθήματα μπορούν να μοντελοποιηθούν τεχνητά στις μηχανές. Συγκεκριμένα και λαμβάνοντας υπόψη τις διακυμάνσεις ανάμεσα στα συναισθήματα στους ανθρώπους ακόμα και κάτω από τις ίδιες συνθήκες, στόχος του **EmBP** αποτέλεσε η βελτιστοποίηση της διαδικασίας μάθησης ενός απλού μοντέλου μηχανικής μάθησης. Τα συναισθήματα που εισάγονται είναι το άγχος (*anxiety*) και η αυτοπεποίθηση (*confidence*). Τα συναισθήματα αυτά θεωρούνται καθοριστικά στην διαδικασία μάθησης και λήψης αποφάσεων στους ανθρώπους. Η υπόθεση που πραγματοποιείται είναι ότι με την εκμάθηση μίας νέας δραστηριότητας τα επίπεδα άγχους είναι υψηλά, ενώ η αυτοπεποίθηση είναι σε χαμηλά επίπεδα. Με εξάσκηση ωστόσο και μετά από την πάροδο κάποιου χρονικού διαστήματος, το άγχος μειώνεται και αυξάνεται η αυτοπεποίθηση. Επομένως, το άγχος και η αυτοπεποίθηση θεωρούνται δύο εξαρτώμενες διαστάσεις, όπου η αυτοπεποίθηση ορίζεται ως ο αρνητικός ρυθμός αλλαγής στο άγχος.



Εικόνα 4.2: Είσοδος/Εξοδος του κρυμμένου νευρώνα στο δίκτυο εφαρμογής του **EmBP**

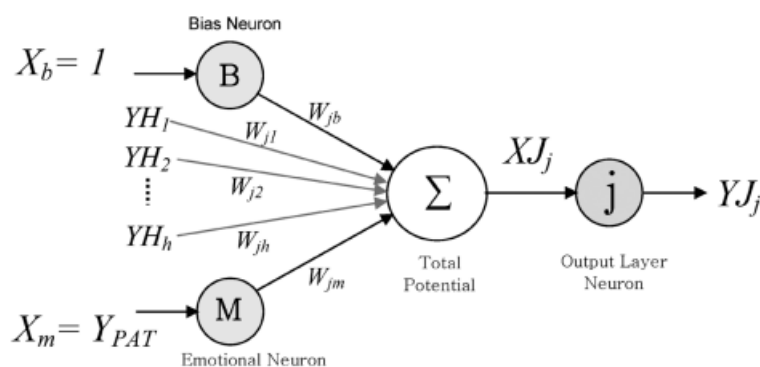
Ο A. Khashman βασιζόμενος στην δουλειά των Rumelhart et al. [32] τροποποιεί την δέλτα ανανέωση των παραμέτρων ενός απλού νευρωνικού δικτύου (*backpropagation*)

προσθέτοντας τις παραμέτρους που περιγράφηκαν παραπάνω. Πιο συγκεκριμένα, η εφαρμογή του αλγορίθμου πραγματοποιείται σε ένα προσθιοδρομικό (feedforward) νευρωνικό δίκτυο με ένα επίπεδο κρυμμένων νευρώνων με σιγμοειδή συνάρτηση ενεργοποίησης. Η είσοδος σε κάθε νευρώνα του κρυμμένου επιπέδου ορίζεται ως το συνολικό δυναμικό (total potential) όλων των δεδομένων εισόδου στον νευρώνα. Το συνολικό αυτό δυναμικό είναι το άθροισμα των γινομένων ανάμεσα στις τιμές εισόδου και των βαρών που έχουν ανατεθεί στις συνάψεις. Πέρα από τα βασικά δυναμικά των δεδομένων εισόδου και της πόλωσης (bias), εισάγεται μία νέα είσοδος που αποτελεί το γενικευμένο μέσο όρο των προτύπων εισόδου. Η τιμή αυτή ορίζεται ως η είσοδος του συναισθηματικού νευρώνα του επιπέδου εισόδου και βασίζεται στην διατύπωση των Baumgartner et al. [51] ότι οι περισσότερες από τις δημοσιευμένες έρευνες νευρολογικής απεικόνισης χρησιμοποιούν οπτικό ερέθισμα προκειμένου να προκαλέσουν την εκδήλωση συναισθημάτων. Αντίστοιχα, ο γενικευμένος αυτός μέσος όρος δίνεται ως είσοδος και στον συναισθηματικό νευρώνα του επιπέδου εξόδου.

Όσον αφορά τις παραμέτρους της συναισθηματικά επαυξημένης οπισθοδιάδοσης, μαζί με τις υπάρχουσες παραμέτρους του ρυθμού μάθησης (learning rate) και του ρυθμού ορμής (momentum) χρησιμοποιούνται και οι παράμετροι του άγχους (anxiety) και της αυτοπεποίθησης (confidence). Οι τελευταίες κινούνται ανάμεσα στο μηδέν και τη μονάδα και ακολουθούν τις παρακάτω παραδοχές:

- Παραδοχή 1: Το επίπεδο του άγχους (anxiety) εξαρτάται από τα πρότυπα εισόδου, όπου τα νέα πρότυπα προκαλούν μεγαλύτερο άγχος. Κατά τη διάρκεια της πρώτης επανάληψης (iteration) η αρχική τιμή του άγχους ορίζεται ίση με 1, δηλαδή ίση με τη μέγιστη δυνατή τιμή.
- Παραδοχή 2: Το επίπεδο άγχους εξαρτάται από το σφάλμα ανάμεσα στην έξοδο του νευρωνικού δικτύου και την επιθυμητή έξοδό του. Αυτού του είδους η ανάδραση χρησιμοποιείται για την αξιολόγηση της απόδοσης του δικτύου. Το άγχος μειώνεται με την ελαχιστοποίηση αυτού του σφάλματος.
- Παραδοχή 3: Το επίπεδο αυτοπεποίθησης αυξάνεται με τη μείωση του επιπέδου του άγχους. Κατά την πρώτη επανάληψη (iteration) η αρχική τιμή της αυτοπεποίθησης ορίζεται στο 0, δηλαδή στην ελάχιστη δυνατή τιμή.

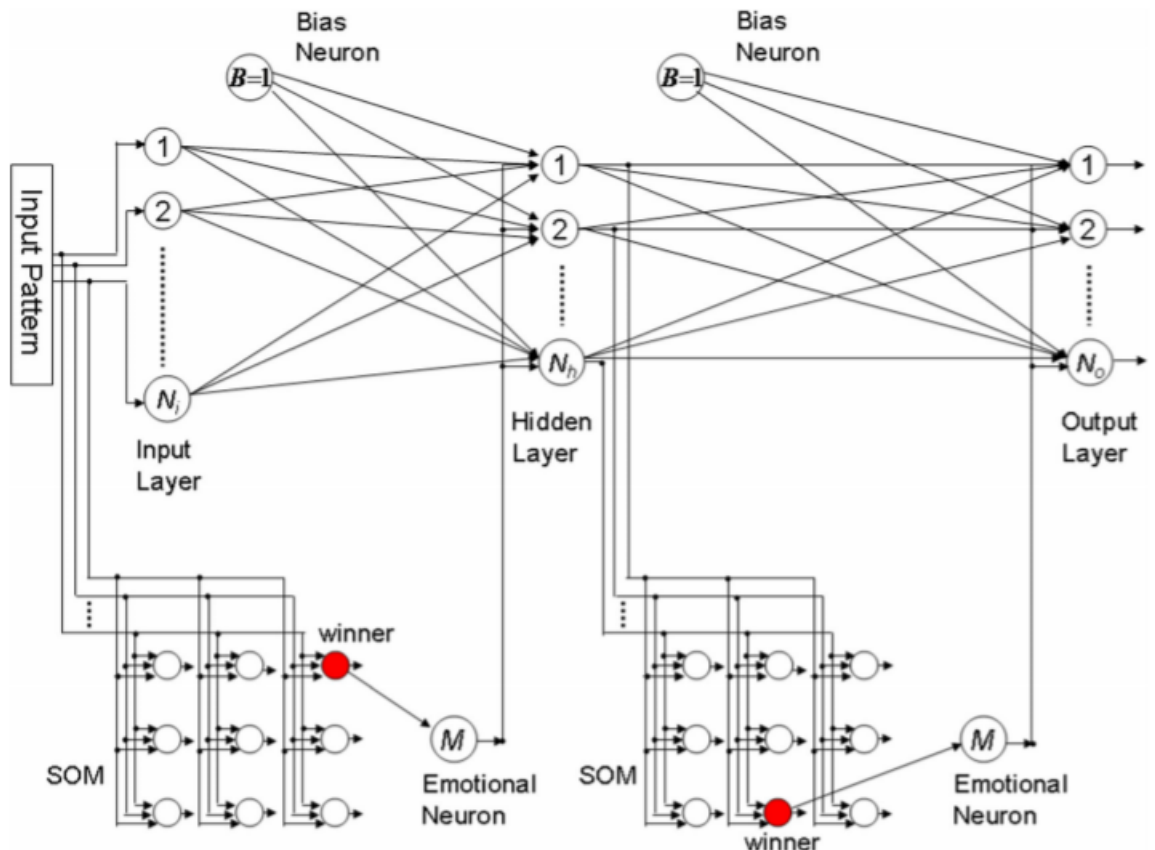
Οι παράμετροι αυτές σε συνδυασμό με τις παραδοχές τους αξιοποιούνται στην ανανέωση των παραμέτρων του δικτύου που αφορούν τη συναισθηματική του είσοδο. Ο τρόπος ανανέωσης τους είναι αντίστοιχος με τον τρόπο ανανέωσης του ρυθμού μάθησης και της ορμής.



Εικόνα 4.3: Είσοδος/Έξοδος του νευρώνα εξόδου στο δίκτυο εφαρμογής του EmBP

4.2.2 Δομή νευρωνικού δικτύου EmSOM

Με βάση τα παραπάνω οι Dai et al. προχώρησαν στην θεωρητική και πειραματική μελέτη του EmBP αλγορίθμου καταλήγοντας στο ότι η ανάθεση της συναισθηματικής εισόδου μάλλον δεν είναι η βέλτιστη. Για παράδειγμα, μπορεί να υπάρχουν δύο εικόνες εισόδου που, παρότι διαφορετικές και διακριτές από ανθρώπινο μάτι, διαθέτουν πολύ όμοιους γενικευμένους μέσους όρους. Σε αυτή την περίπτωση το δίκτυο δεν καταφέρνει να τις διακρίνει και τις θεωρεί παραπλήσιες. Ακόμα, θεωρείται υπεραπλουστευμένη η τροφοδότηση των συναισθηματικών νευρώνων του κρυμμένου επιπέδου και του επιπέδου εξόδου με την ίδια τιμή. Για το λόγο αυτό αντικαθιστούν το συναισθηματικό νευρώνα εισόδου με ένα χάρτη αυτο-οργάνωσης. Η είσοδος των συναισθηματικών νευρώνων επομένως δεν είναι απλά ο γενικευμένος μέσος όρος των προτύπων εισόδου, αλλά προκύπτει από τον νικητή νευρώνα του αντίστοιχου χάρτη αυτο-οργάνωσης. Πιο συγκεκριμένα, πρώτα ένα SOM δίκτυο που συνδέεται με το κρυμμένο επίπεδο του δικτύου εκπαιδεύεται πάνω σε όλα τα δεδομένα εισόδου. Στη συνέχεια, για κάποιο συγκεκριμένο πρότυπο εισόδου λαμβάνεται ο μέσος όρος ως προς τη διάστασή του του διανύματος βαρών του νικητή νευρώνα του SOM δικτύου και χρησιμοποιείται ως είσοδος στο συναισθηματικό νευρώνα του κρυμμένου επιπέδου. Αντίστοιχα, λαμβάνοντας όλες τις εξόδους από το κρυμμένο επίπεδο του νευρωνικού δικτύου εκπαιδεύεται το SOM δίκτυο που συσχετίζεται με το επίπεδο εξόδου. Ακολούθως, για κάθε έξοδο του κρυμμένου επιπέδου υπολογίζεται ο μέσος όρος των βαρών του νικητή νευρώνα του αντίστοιχου SOM δικτύου και τροφοδοτείται στον συναισθηματικό νευρώνα του επιπέδου εξόδου.



Εικόνα 4.4: EmSOM τοπολογία

Οι λόγοι που οδήγησαν στην υιοθέτηση των χαρτών αυτο-οργάνωσης για την είσοδο των συναισθηματικών νευρώνων παρουσιάζονται συνοπτικά παρακάτω:

- Καταρχάς, πρότυπα εισόδου που ανήκουν σε διαφορετικές κατηγορίες αναδεικνύουν, το πιο πιθανό, διαφορετικούς νικητές νευρώνες στα SOM δίκτυα του κρυμμένου επιπέδου και του επιπέδου εξόδου. Με τον τρόπο αυτό οι τιμές εισόδου στους συναισθηματικούς νευρώνες έχουν διαφορετικές τιμές και πιθανά το γεγονός αυτό να βελτιώνει τον ρυθμό αναγνώρισης.
- Λόγω των ιδιοτήτων τους τα SOM δίκτυα που εισάγονται στο νευρωνικό δίκτυο παρέχουν μία αναπαράσταση του χώρου των δεδομένων εισόδου και των εξόδων του κρυμμένου επιπέδου. Επομένως, η γενική εποπτεία των δεδομένων που επιχειρείται να αποτυπωθεί στην συναισθηματική είσοδο βελτιστοποιείται. Ακόμα, τοπολογικά στα SOM δίκτυα απεικονίζονται τα χαρακτηριστικά (**features**) των δεδομένων ως συστάδες νευρώνων.
- Ο χάρτης χαρακτηριστικών (**feature map**) που προκύπτει αντικατοπτρίζει τη φύση της κατανομής των δεδομένων εισόδου. Σαφέστερα, περιοχές στον χώρο εισόδου όπου τα δείγματα λαμβάνονται με μεγάλη πιθανότητα αντιστοιχίζονται σε μεγαλύτερες περιοχές και στον χώρο εξόδου. Με αυτόν τον τρόπο επιτυγχάνεται καλύτερη ευκρίνεια σε σύγκριση με περιοχές του χώρου εισόδου που δεν δειγματοληπτούνται με την ίδια πιθανότητα.
- Όσον αφορά τον χώρο εισόδου, ακόμα και σε μη γραμμικές κατανομές, οι χάρτες αυτο-οργάνωσης έχουν τη δυνατότητα να επιλέξουν το βέλτιστο σύνολο χαρακτηριστικών για την προσομοίωση της κατανομής εισόδου.

4.2.2.1 Κρυφό επίπεδο νευρώνων του EmSOM

Οι Dai et al. επιλέγουν σε αντιστοιχία με τον A. Khashman ένα προσθιοδρομικό νευρωνικό δίκτυο (**feed-forward neural network**) με μοναδικό επίπεδο κρυμμένων νευρώνων. Ωστόσο, σημειώνουν ότι η δομή αυτή μπορεί να γενικευτεί και σε περισσότερα κρυμμένα επίπεδα νευρώνων. Σε αντιστοιχία με τους Rumelhart et al. και Khashman [32], [27] επιλέγουν την σιγμοειδή συνάρτηση ως συνάρτηση ενεργοποίησης των νευρώνων του κρυμμένου επιπέδου. Συνεπώς, η έξοδος κάθε κρυμμένου νευρώνα σε αυτό το επίπεδο στην επανάληψη (**iteration**) n υπολογίζεται ως εξής:

$$y_j(n) = \phi(v_j(n)) = \frac{1}{1 + \exp(-v_j(n))}$$

όπου $v_h(n)$ και $y_j(n)$ αποτελούν τις τιμές εισόδου και εξόδου του j -οστού νευρώνα σε αυτό το επίπεδο.

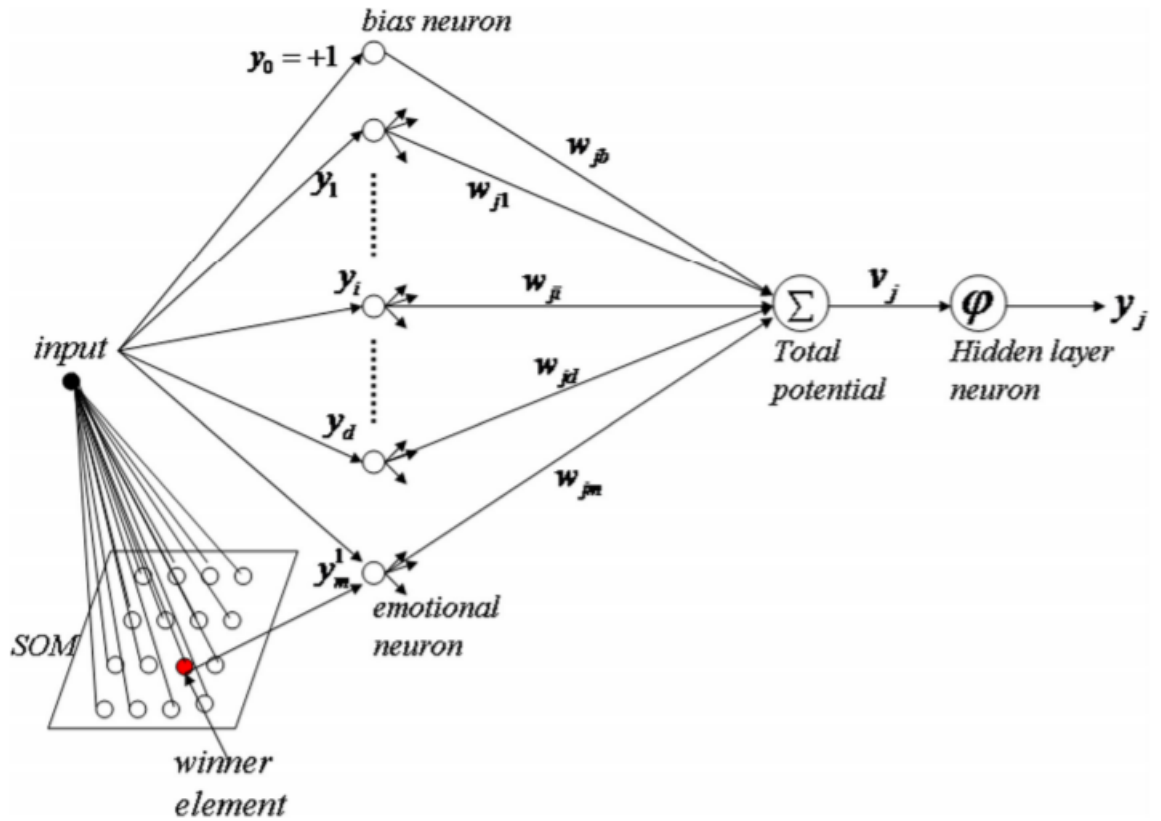
Η είσοδος σε κάθε κρυμμένο νευρώνα υπολογίζεται ως εξής:

$$v_j(n) = v_{jc}(n) + v_{jb}(n) + v_{jm}(n)$$

, όπου $v_{jc}(n)$ αποτελεί τη συμβατική τιμή εισόδου σε κάθε κρυμμένο νευρώνα βάσει του αλγορίθμου οπισθοδιάδοσης (**backpropagation**) και ορίζεται ως εξής:

$$v_{jc}(n) = \sum_{i=1}^d w_{ji}(n) y_i(n)$$

όπου $w_{ji}(n)$ αναπαριστά την τιμή του βάρους που ενώνει τον νευρώνα εισόδου i με τον κρυμμένο νευρώνα j στην επανάληψη (**iteration**) n , $y_j(n)$ αναπαριστά την i -οστή είσοδο



Εικόνα 4.5: Ροή πληροφορίας για το κρυφό επίπεδο νευρώνων.

του δικτύου στην επανάληψη (iteration) n , και d αποτελεί η διάσταση των δεδομένων εισόδου. Το $v_{jb}(n)$ υπολογίζεται ως:

$$v_{jb}(n) = w_{jb}(n)y_0$$

και

$$y_0 = 1$$

Ο συμβολισμός $w_{jb}(n)$ αναπαριστά τα βάρη των ενώσεων ανάμεσα στον νευρώνα πόλωσης (bias) του κρυμμένου επιπέδου νευρώνων και τον j -στό κρυμμένο νευρώνα, ενώ y_0 είναι η είσοδος του νευρώνα αυτού.

Για τον υπολογισμό του νικητή νευρώνα του χάρτη αυτο-οργάνωσης του κρυμμένου επιπέδου το δίκτυο τροφοδοτείται με όλα τα δεδομένα εισόδου. Μετά από έναν αριθμό από εποχές εκπαίδευσης το SOM δίκτυο επιτυγχάνει σε κάποιο βαθμό αυτο-οργάνωση. Τότε επιλέγεται ένα στοιχείο νικητής i^* ως το πιο κοντινό στοιχείο στο πρότυπο εισόδου. Επομένως στο EmSOM το $v_{jm}(n)$ ορίζεται ως εξής:

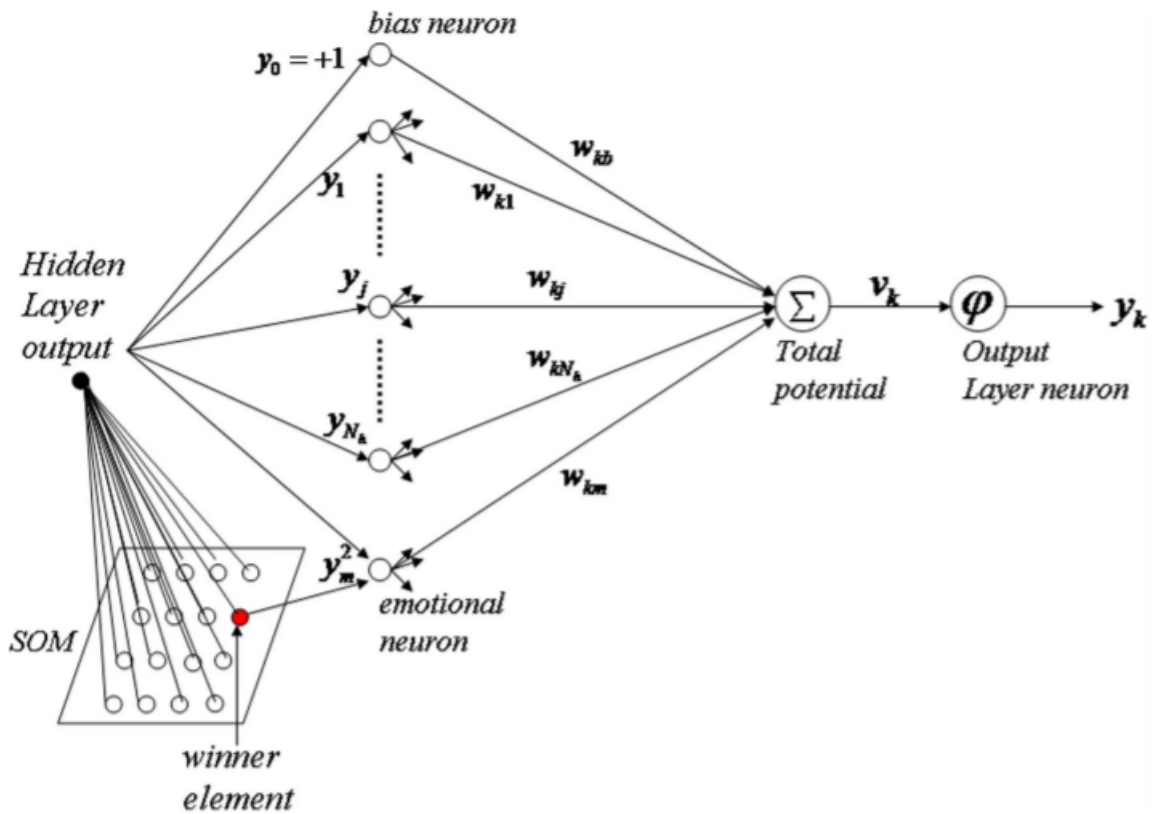
$$v_{jm}(n) = w_{jm}(n)y_m^1(n)$$

και

$$y_m^1(n) = \frac{\sum_{i=1}^d w_{i^*,i}^1(n)}{d}$$

4.2.2.2 Επίπεδο νευρώνων εξόδου του EmSOM

Όπως φαίνεται και στην Εικόνα 4.6 που απεικονίζει την λεπτομερή ροή σήματος στον νευρώνα εξόδου k , και εδώ υιοθετείται η σιγμοειδής συνάρτηση ως συνάρτηση ενεργοποίησης αυτού του επιπέδου.



Εικόνα 4.6: Ροή πληροφορίας για το επίπεδο νευρώνων εξόδου.

Η έξοδος κάθε νευρώνα εξόδου υπολογίζεται ως εξής:

$$y_k(n) = \phi(v_k(n)) = \frac{1}{1 + \exp(-v_k(n))}$$

, όπου οι $v_k(n)$ και $y_k(n)$ αποτελούν τις τιμές εισόδου και εξόδου του k -οστού νευρώνα στο επίπεδο εξόδου στην n -οστή επανάληψη (iteration).

Πανομοιότυπα με το κρυμμένο επίπεδο, το $v_k(b)$ υπολογίζεται:

$$v_k(n) = v_{kc}(n) + v_{kb}(n) + v_{km}(n)$$

και αναπαριστά την συμβατική τιμή εισόδου του k -οστού νευρώνα στην n -οστή επανάληψη (iteration). Επομένως το $v_{kc}(n)$:

$$v_{kc}(n) = \sum_{j=1}^{N_h} w_{kj}(n) y_j(n)$$

όπου $w_{kj}(n)$ αναπαριστά το βάρος της σύνδεσης ανάμεσα στον j -οστό κρυμμένο νευρώνα με το k -οστό νευρώνα εξόδου, και $y_j(n)$ είναι η έξοδος από το j -οστό κρυμμένο νευρώνα στην επανάληψη (iteration) n . N_h είναι το μέγεθος του κρυμμένου επιπέδου.

Η ποσότητα $v_{kb}(n)$ υπολογίζεται από τον νευρώνα πόλωσης του επιπέδου εξόδου και τα σχετικά βάρη.

$$v_{kb}(n) = w_{kb}(n)y_0$$

και

$$y_0 = 1$$

,όπου $w_{kb}(n)$ αναπαριστά τα βάρη των ενώσεων ανάμεσα στον νευρώνα πόλωσης (bias) του επιπέδου εξόδου νευρώνων και τον k -οστό νευρώνα εξόδου, ενώ y_0 είναι η είσοδος του νευρώνα αυτού.

Αντίστοιχα για το SOM δίκτυο του επιπέδου εξόδου, ο υπολογισμός της εισόδου του συναισθηματικού νευρώνα υπολογίζεται ως εξής:

$$v_{km}(n) = w_{km}(n)y_m^2(n)$$

και

$$y_m^2(n) = \frac{\sum_{i=1}^N h w_{i*,i}^2(n)}{N_h}$$

Για τον υπολογισμό του $y_m^2(n)$ ακολουθείται η ίδια διαδικασία όπως και με τον χάρτη αυτο-οργάνωσης του κρυμμένου επιπέδου. Το εν λόγω δίκτυο εκπαιδεύεται σε όλες τις εξόδους του κρυμμένου επιπέδου και επιτυγχάνει μετά από κάποιες εποχές εκπαίδευσης έναν βαθμό αυτο-οργάνωσης. Ο νευρώνας νικητής i^* επιλέγεται ανάλογα την εκάστοτε έξοδο του κρυμμένου επιπέδου και η ποσότητα $w_{i*,i}^2(n)$ αντιστοιχεί στο διάνυσμα βαρών του στην n -οστή επανάληψη (iteration).

4.2.3 Συναισθηματική οπισθοδιάδοση (EmBP)

Εδώ οι Dai et al.[9] συμφωνούν στα περισσότερα σημεία με τον αλγόριθμο του Khashman. Η λογική παραμένει η ίδια, ότι δηλαδή στην αρχή της εκπαίδευσης το άγχος είναι στα υψηλότερα επίπεδα, ενώ η αυτοπεποίθηση στα χαμηλότερα επίπεδα. Ακόμα, το άγχος εξαρτάται από το σφάλμα ανάμεσα στην έξοδο του δικτύου και την επιθυμητή έξοδο και μεταβάλλεται με την πάροδο των εποχών εκπαίδευσης. Επομένως, με την μείωση του σφάλματος του δικτύου μειώνονται και τα επίπεδα άγχους. Αντίστοιχα, λόγω της εξάρτησης ανάμεσα στο άγχος και την αυτοπεποίθηση, με την πάροδο των εποχών εκπαίδευσης η αυτοπεποίθηση αυξάνεται.

Ωστόσο εδώ υπάρχουν διαφοροποιήσεις από τον αρχικό αλγόριθμο. Οι δύο κύριες διαφορές στους υπολογισμούς αυτούς αποτελούν οι αποκλίσεις στις αρχικοποιήσεις των συναισθηματικών παραμέτρων, αλλά και στον τρόπο που ανανεώνονται σε κάθε εκπαίδευση, έτσι ώστε να υποστηρίζεται η online εκπαίδευση και όχι η batch, όπως στον βασικό αλγόριθμο. Οι τροποποιημένες αυτές μεταβολές των συναισθηματικών παραμέτρων απεικονίζονται στις εξισώσεις που ακολουθούν.

4.2.3.1 Υπολογισμός συναισθηματικών παραμέτρων

Η τιμή της παραμέτρου του άγχους αρχικοποιείται ως εξής:

$$axt(0) = Y_{AllPAT} + E(0)$$

, όπου το Y_{AllPAT} αναπαριστά το γενικευμένο μέσο όρο όλων των προτύπων εισόδου σε κάθε εποχή εκπαίδευσης και υπολογίζεται:

$$Y_{AllPAT} = \frac{\sum_{p=1}^{N_p} Y_{PAT}}{N_p}$$

Το N_p είναι το μέγεθος του συνόλου δεδομένων εκπαίδευσης και Y_{PAT} η μέση τιμή της εκάστοτε εικόνας εισόδου $P(x, y)$. Δηλαδή:

$$Y_{PAT} = \frac{\sum_{x=1, y=1}^{x_{max}, y_{max}} P(x, y)}{x_{max} y_{max}}$$

Το $E(0)$ αποτελεί την αρχικοποίηση σφάλματος και υπολογίζεται ως εξής:

$$E(0) = \sum_{p=1}^{N_p} \frac{\sqrt{\sum_{k=1}^{N_0} (t_{pk}(0) - y_{pk}(0))^2}}{N_p}$$

, όπου $(t_{pk}(0) - y_{pk}(0))$ σηματοδοτεί το σήμα σφάλματος ανάμεσα στο k -οστό νευρώνα εξόδου για το p -οστό πρότυπο εκπαίδευσης.

Η τιμή του άγχους ανανεώνεται σε κάθε επανάληψη n ως εξής:

$$axt(n) = Y_{AllPAT} + \frac{E(n)}{N_0} \quad (4.3.2.1.1)$$

και

$$E(n) = \frac{1}{2} \sum_{k=1}^{N_0} (t_k(n) - y_k(n))^2 \quad (4.3.2.1.2)$$

, όπου $t_k(n)$ και $y_k(n)$ συνιστούν την έξοδο του δικτύου και την επιθυμητή έξοδο αντίστοιχα για τον k -οστό νευρώνα εξόδου στην επανάληψη n . Το N_0 αναπαριστά τον αριθμό των νευρώνων στο επίπεδο εξόδου. Βάσει της τιμής του άγχους η τιμή της αυτοπεποίθησης ορίζεται ως εξής:

$$cfd(n) = atx(n) - axt(0) \quad (4.3.2.1.3)$$

όπου $axt(0)$ η αρχική τιμή άγχους και $axt(n)$ η τιμή του άγχους στην n -οστή επανάληψη.

4.2.3.2 Ανανέωση παραμέτρων δικτύου κατά την εκπαίδευση

Δεδομένης της επιθυμητής τιμής εξόδου $t_k(n)$ του k -οστού νευρώνα εξόδου στην n -οστή επανάληψη και της αντίστοιχης πραγματικής τιμής εξόδου $y_k(n)$, η συνάρτηση σφάλματος ορίζεται το άθροισμα ελάχιστων τετραγώνων και αναπαρίσταται παραπάνω με την εξίσωση 4.3.2.1.2.

Όπως και στο EmBP αλλά και στον κλασικό αλγόριθμο **backpropagation**, το σήμα σφάλματος διαδίδεται προς την αρχή του δικτύου έτσι ώστε να προσαρμοστούν όλες οι παράμετροί του στο τέλος κάθε επανάληψης εκπαίδευσης (**training iteration**). Τα βάρη ανάμεσα στο επίπεδο εξόδου και το κρυμμένο επίπεδο ανανεώνονται βάσει της παρακάτω εξίσωσης:

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{c_{kj}(n)} + \Delta w_{m_{kj}(n)}$$

όπου $\Delta w_{c_{kj}}(n)$ είναι η προσαρμογή βάρους ανάμεσα στον j -οστό κρυμμένο νευρώνα και τον k -οστό νευρώνα εξόδου βάσει του συμβατικού **backpropagation** αλγορίθμου, και

$\Delta w_{m_{kj}}(n)$ αναπαριστά την προσαρμογή βαρών όσον αφορά την συναισθηματική πληροφορία. Το $\Delta w_{c_{kj}}(n)$ υπολογίζεται ως εξής:

$$\Delta W_{c_{kj}}(n) = \eta \delta_k(n) y_j(n) + \alpha [\Delta w_{c_{kj}}(n-1)]$$

Το $\delta_k(n)$ ορίζεται ως το σήμα σφάλματος και αναπαρίσταται ως εξής:

$$\delta_k(n) = \phi'_k(v_k(n))(t_k(n) - y_k(n))$$

$$\delta_k(n) = y_k(n)(1 - y_k(n))(t_k(n) - y_k(n))$$

όπου η είναι ο ρυθμός μάθησης, α η ορμή, και $\Delta w_{c_{kj}}(n-1)$ η προηγούμενη αλλαγή στα βάρη. Η ανανέωση των βαρών με την συναισθηματική έννοια για τον EmSOM αλγόριθμο ορίζεται ως εξής:

$$\Delta w_{m_{kj}}(n) = axt(n) \delta_k(n) y_m^2(n) + cfd(n) [\Delta w_{c_{kj}}(n-1)]$$

, όπου $axt(n)$ και $cfd(n)$ αποτελούν τις τιμές άγχους και αυτοπεποίθησης στην επανάληψη (iteration) n αντίστοιχα. Ο συμβολισμός $y_m^2(v)$ αναπαριστά τη μέση τιμή των βαρών ως προς τη διάσταση του διανύσματος του νικητή νευρώνα i^* στο SOM δίκτυο που σχετίζεται με το επίπεδο εξόδου του δικτύου. Αντίστοιχα για τα βάρη ανάμεσα στο κρυμμένο επίπεδο και το επίπεδο εισόδου, οι παράμετροι αυτοί ανανεώνονται βάσει της ακόλουθης εξίσωσης:

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{c_{ji}}(n) + \Delta w_{m_{ji}}(n)$$

, με το $\Delta w_{c_{ji}}(n)$ να αποτελεί την προσαρμογή βάρους της σύνδεσης ανάμεσα στον i -οστό νευρώνα εισόδου και τον j -οστό κρυμμένο νευρώνα με τη συμβατική έννοια, ενώ το $\Delta w_{m_{ji}}(n)$ με τη συναισθηματική έννοια. Το $\Delta w_{c_{ji}}(n)$ υπολογίζεται:

$$\Delta W_{c_{ji}}(n) = \eta \delta_j(n) y_i(n) + \alpha [\Delta w_{c_{ji}}(n-1)]$$

, όπου $\delta_j(n)$ το αντίστοιχο σήμα σφάλματος που ορίζεται ως εξής:

$$\delta_j(n) = y_j(n)(1 - y_j(n)) \sum_{k=1}^{N_o} w_{kj}(n) \delta_k(n)$$

Η ποσότητα $\Delta w_{m_{ji}}(n)$ ορίζεται ως εξής:

$$\Delta w_{m_{ji}}(n) = axt(n) \delta_j(n) y_m^1(n) + cfd(n) [\Delta w_{c_{ji}}(n-1)]$$

, με το $y_m^1(n)$ να αναπαριστά τη μέση τιμή ως προς το μήκος του διανύσματος των βαρών του νευρώνα νικητή i^* του δικτύου SOM που σχετίζεται με το κρυμμένο επίπεδο νευρώνων.

Κεφάλαιο 5

Υλοποίηση και αποτελέσματα

Η υλοποίηση του υβριδικού δικτύου που παρουσιάστηκε στο Κεφάλαιο 4 επιλέχθηκε να πραγματοποιηθεί σε γλώσσα `python`. Δεδομένης της μεγάλης ποικιλίας εργαλείων και υπολογιστικών πλαισίων (`frameworks`) στη προγραμματιστική αυτή γλώσσα, παρέχεται ευχέρεια για γρήγορη και αποδοτική υλοποίηση ετερόκλητων νευρωνικών δικτύων. Ο κώδικας που υλοποιήθηκε και αναλύεται στις ακόλουθες ενότητες μπορεί να βρεθεί στο `GitHub`¹.

5.1 Επιλογή εργαλείων υλοποίησης

Ανάμεσα στα διαθέσιμα υπολογιστικά εργαλεία επιλέχθηκε το `PyTorch framework` για την υλοποίηση του `feed-forward` μέρους του νευρωνικού δικτύου. Η επιλογή αυτή βασίζεται στην ευκολία με την οποία μπορούν να δημιουργηθούν `feed-forward` δίκτυα με πολύ μεγάλο βαθμό παραμετροποίησης στο `PyTorch`, καθώς και η παροχή έτοιμων εργαλείων και δομών, όπως ο δυναμικός γράφος `autograd` που παρέχει αυτόματη διαφοροποίηση των παραμέτρων του δικτύου ως προς το σφάλμα πρόβλεψης.

Για την υλοποίηση του `SOM` δικτύου επιλέχθηκε η βιβλιοθήκη `quicksom`². Με τη χρήση αυτής της βιβλιοθήκης μπορεί να δημιουργηθεί ένα `SOM` αντικείμενο σε οποιοδήποτε μέγεθος (`grid size`) και με την επιλογή περιοδικής τοπολογίας. Ακόμα, δίνονται βαθμοί ελευθερίας στην επιλογή του πλήθους των εποχών εκπαίδευσης, του μεγέθους `batch`, των ενδιάμεσων εκτυπώσεων κ.ο.κ. Διαθέσιμα είναι επίσης και τρία έτοιμα `scripts` για την εκπαίδευση ενός `SOM`, την προαιρετική δημιουργία συστάδων (`clusters`) με μία διεπαφή γραφικού περιβάλλοντος (`Graphical User Interface, GUI`) και την πρόβλεψη επιρροής στις συστάδες με την εισαγωγή νέων δεδομένων. Δεδομένου ότι η υλοποίηση αυτής της βιβλιοθήκης είναι ανοιχτού κώδικα (`open-source`), η τροποποίηση βασικών λειτουργικότητας των `SOM` δικτύων οδήγησε στην υλοποίηση τους όπως ακριβώς περιγράφονται στην ερευνητική εργασία των `Dai et al.` Ακόμα, με αυτή την βιβλιοθήκη παρέχεται η δυνατότητα πραγματοποίησης υπολογισμών σε κάρτα γραφικών (`Graphics Processing Unit, GPU`) για την επίτευξη πιο γρήγορων υπολογισμών.

5.2 Υλοποίηση

Η υλοποίηση του υβριδικού νευρωνικού δικτύου εφαρμόστηκε σε δύο σύνολα δεδομένων (`datasets`). Αρχικά, εφαρμόστηκε στο ίδιο σύνολο δεδομένων αναγνώρισης προσώπου που περιγράφεται και στην ερευνητική εργασία παρουσίασης του `EmSOM` δικτύου [9]. Στη συνέχεια διερευνήθηκε η απόδοσή του σε διαφορετικό σύνολο δεδομένων που περιλαμβάνει εικόνες εκφράσεων προσώπων. Για το λόγο αυτό, οι υλοποιήσεις του δικτύου στην περίπτωση κάθε συνόλου δεδομένων διαφοροποιούνται στις παραμετροποιήσεις τους.

¹https://github.com/akarkoglou/DSML_thesis_hybrid_sentiment

²<https://github.com/bougui505/quicksom>

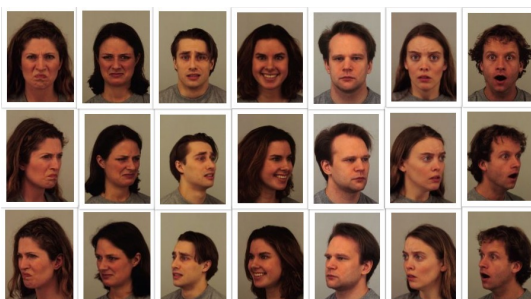
5.2.1 Σύνολα Δεδομένων

Το σύνολο δεδομένων που περιγράφεται στην παρουσίαση του επιλεγμένου νευρωνικού δικτύου είναι το ORL (Our Database of Faces) των AT&T εργαστηρίων [34]. Το dataset αυτό περιλαμβάνει 400 εικόνες από 40 διαφορετικά πρόσωπα. Για κάποια απο αυτά οι φωτογραφίες έχουν ληφθεί σε διαφορετικές χρονικές στιγμές, με διαφορετικό φωτισμό, εκφράσεις προσώπου (ανοιχτά/κλειστά μάτια, με χαμόγελο/χωρίς) και με διαφορετικές λεπτομέρειες προσώπου (όπως η ύπαρξη ή όχι γυαλιών μυωπίας). Όλες οι εικόνες έχουν ληφθεί με ένα σκουρόχρωμο και ομογενές φόντο και με τα πρόσωπα σε μία μετωπική στάση με περιθώριο μικρής πλάγιας κίνησης. Τέλος, το μέγεθος των εικόνων είναι 92x112 pixels με 256 διαβαθμίσεις γκρι ανά pixel.



Εικόνα 5.1: Εικόνες από το ORL σύνολο δεδομένων

Το σύνολο δεδομένων που επιλέγεται να διερευνηθεί η απόδοση του υβριδικού Em-SOM είναι το KDEF (The Karolinska Directed Emotional Faces) [15]. Αυτό το σύνολο δεδομένων περιλαμβάνει 4900 εικόνες από εκφράσεις προσώπων. Οι εικόνες αυτές απεικονίζουν 70 διαφορετικούς ανθρώπους που παίρνουν 7 διαφορετικές συναισθηματικές εκφράσεις. Κάθε τέτοια έκφραση απαθανατίζεται από 5 διαφορετικές οπτικές γωνίες. Τα πρόσωπα που συμμετέχουν δεν έχουν χαρακτηριστικά προσώπου, όπως μούσι, μουστάκι, σκουλαρίκια ή γυαλιά μυωπίας. Ακόμη, όλοι φορούν την ίδια γκρι μπλούζα και στέκονταν σε 3 μέτρα απόσταση από το φακό της κάμερας. Το μέγεθος των εικόνων είναι ίσο με 562x762 pixels.



Εικόνα 5.2: Εικόνες από το KDEF σύνολο δεδομένων

5.2.1.1 Προ-επεξεργασία δεδομένων

Και για τα δύο σύνολα δεδομένων αξιοποιείται ο PyTorch dataloader ImageFolder, ο οποίος φορτώνει τα δεδομένα που βρίσκονται σε δενδρική μορφή στο σύστημα αρχείων.

Σύμφωνα με αυτή την δομή όλες οι εικόνες που αντιστοιχούν σε μία κατηγορία προς ταξινόμηση βρίσκονται κάτω από τον ίδιο φάκελο με το όνομα της κλάσης. Με τη φόρτωση των δεδομένων εφαρμόζονται σε αυτά μετασχηματισμοί προκειμένου οι εικόνες να έρθουν σε μορφή αξιοποιήσιμη από το νευρωνικό δίκτυο.

Στην περίπτωση του **ORL dataset** εφαρμόζονται οι προδιαγραφές που διατυπώθηκαν ως βέλτιστες με σκοπό την βελτίωση της ορθότητας (**accuracy**) του δικτύου. Οι εικόνες προσαρμόζονται σε μέγεθος **46x56 pixels** και οι τιμές των **pixels** κανονικοποιούνται στο εύρος 0 έως 0.1. Η κανονικοποίηση αυτή πραγματοποιείται για τον περιορισμό του κορεσμού του δικτύου. Το σύνολο δεδομένων χωρίζεται σε διάφορα ποσοστά για τη δημιουργία των συνόλων εκπαίδευσης και αξιολόγησης, και οι εικόνες κάθε συνόλου επιλέγονται τυχαία από το γενικό σύνολο. Το μέγεθος του **batch** με το οποίο τροφοδοτούνται τα δεδομένα στο δίκτυο ορίζεται στις 10 εικόνες.

Αντίστοιχα και το **KDEF dataset** προσαρμόζεται στα **46x56 pixels** με την συνάρτηση **resize()** των **PyTorch** μετασχηματισμών. Λόγω του μεγάλου παράγοντα σμίξυσης ωστόσο, πολλά χαρακτηριστικά και πρότυπα των δεδομένων εισόδου χάνονται. Η διατήρηση μεγάλου μεγέθους εικόνων θα αύξανε σημαντικά τον αριθμό των παραμέτρων δικτύου, με αποτέλεσμα και την αναπόφευκτη αύξηση της διακύμανσής του (**variance**). Λόγω της ευαισθησίας του δικτύου στην υπερεκπαίδευση, όπως θα αναλυθεί και στη συνέχεια, το μέγεθος του **batch** μειώνεται και ορίζεται σε 1, δηλαδή **online** μάθηση. Το μικρό μέγεθος **batch** αναμένεται να προσφέρει κανονικοποίηση (**regularization**) ως κάποιο βαθμό, αφού αυτός ο τρόπος τροφοδότησης δικτύου εισάγει θόρυβο και περιορίζει το δίκτυο από το να υπερχεπαιδεύεται στα δεδομένα. Τέλος, όπως και στο προηγούμενο σύνολο δεδομένων, οι τιμές των **pixels** κανονικοποιούνται στο διάστημα 0 έως 0.1, ώστε να περιοριστεί ο βαθμός κορεσμού του δικτύου.

5.2.2 Βασική Δομή Δικτύου

Παρά το γεγονός ότι η δομή του νευρωνικού δικτύου **EmSOM** σε καθεμία από τις περιπτώσεις είναι η ίδια στη βάση του, η παραμετροποίησή τους διαφέρει. Για την υλοποίηση του δικτύου δημιουργείται η κλάση **EmSOM** που επεκτείνει την **PyTorch** κλάση **nn.Module**, η οποία αποτελεί την βάση για τα νευρωνικά δίκτυα στο **framework** αυτό. Το δίκτυο αρχικοποιείται με δύο γραμμικά επίπεδα, και ως συνάρτηση ενεργοποίησης δικτύου ορίζεται η σιγμοειδής.

Ακόμα η συνάρτηση **forward()** του **nn.Module** που επεκτείνει (**extension**) το **EmSOM** δίκτυο, ορίζεται έτσι ώστε να υπολογίζει για κάθε δεδομένο εισόδου το μέτρο των βαρών του νικητή νευρώνα και να το προσαρτά στο τέλος της συμβατικής εισόδου. Με τον τρόπο αυτό διατηρείται η γραμμικότητα των εξισώσεων που περιγράφουν την συμβατική και την συναισθηματική ροή πληροφορίας.

Για τον υπολογισμό της συναισθηματικής εισόδου για το κρυμμένο, αλλά και για το επίπεδο εξόδου δημιουργείται μία νέα συνάρτηση που ονομάζεται **find_bmus()**. Η συνάρτηση αυτή υλοποιεί τη διαδικασία ανεύρεσης του νικητή νευρώνα του αντίστοιχου χάρτη αυτο-οργάνωσης και τον υπολογισμό του μέσου όρου ως προς το μέγεθος του του διανύσματος των βαρών του.

5.2.3 Υλοποίηση δικτύου **SOM**

Για την υλοποίηση του δικτύου **SOM** χρησιμοποιείται η βιβλιοθήκη **quicksom** λόγω της δυνατότητας τροποποίησης του κώδικα υλοποίησης βάσει των τιμών των παραμέτρων που βρέθηκαν να λειτουργούν βέλτιστα για την απόδοση του δικτύου. Πιο συγκεκριμένα, τροποποιείται η συνάρτηση **scheduler()** που καθορίζει τον ρυθμό ανανέωσης του ρυθμού μάθησης, έτσι ώστε ο γραμμικός τρόπος υπολογισμού του αντίστοιχου βήματος να συνάδει

με την περιγραφή των Dai et al. Το βήμα ανανέωσης του ρυθμού μάθησης υπολογίζεται βάσει του παρακάτω τύπου:

$$\frac{a}{t + b}$$

όπου

$$b = \frac{\text{training_length} - 1}{100 - 1}$$

και

$$a = b * \text{initial_learning_rate}$$

Η τιμή του αριθμού των επαναλήψεων εκπαίδευσης ορίζεται ίση με 100, ενώ ο αρχικός ρυθμός μάθησης ορίζεται σε 0.5 και 0.05 για την περίοδο ακριβούς προσαρμογής (*fine-tuning*). Επιπλέον, το βήμα ανανέωσης της ακτίνας της συνάρτησης γειτνίασης (εδώ επιλέγεται η συνάρτηση Gauss) υπολογίζεται ως εξής:

$$\frac{\text{initial_radius} * \text{final_radius}}{\text{training_length} - 1}$$

Στην περίπτωση του ORL συνόλου δεδομένων, η αρχική τιμή της ακτίνας της συνάρτησης γειτνίασης Gauss ορίζεται ίση με 5, ενώ η τελική τιμή ίση με 1.

Όμοια με τους Dai et al., μέσω δοκιμών προσδιορίζονται και οι παράμετροι για τους χάρτες αυτο-οργάνωσης στην περίπτωση του KDEP συνόλου δεδομένων, έτσι ώστε οι δομές αυτές να συγκλίνουν σε ορισμένο βαθμό και να επιτυγχάνουν κάποιο επίπεδο αυτο-οργάνωσης. Το μέγεθος του δικτύου SOM που σχετίζεται με το χτυπημένο επίπεδο αυξάνεται λόγω της αύξησης και των δεδομένων σε αυτό το *dataset*. Στο ίδιο SOM οι τιμές του αρχικού και τελικού ρυθμού μάθησης και της ακτίνας της συνάρτησης γειτνίασης παραμένουν οι ίδιες. Ωστόσο στο δίκτυο SOM που συνδέεται με το επίπεδο εξόδου ο αρχικός ρυθμός μάθησης ορίζεται μικρότερος και ίσος με 0.5, καθώς παρατηρούνται μεγάλες διακυμάνσεις στη σύγκλιση της απώλειας για μεγαλύτερες τιμές.

Όσον αφορά την εκτέλεση των υπολογισμών σχετικά με τους χάρτες αυτο-οργάνωσης, εδώ αξιοποιείται η κάρτα γραφικών (GPU) αφού δίνεται ως επιλογή από τη βιβλιοθήκη για γρηγορότερους υπολογισμούς.

5.2.4 Υλοποίηση του Emotional Backpropagation (EmBP)

Η υλοποίηση του αλγορίθμου στοχαστικής κλίσης καθόδου (*stochastic gradient descent, SGD*) στο PyTorch (`torch.optim.SGD`) βασίζεται στις παρακάτω εξισώσεις:

$$u = \rho * u + g$$

$$p = p - lr * u$$

, όπου ρ η ορμή (*momentum*), u η ταχύτητα, και g το ανάδελτα (*gradient*) της παραμέτρου p . Επομένως ο ρυθμός μάθησης lr εφαρμόζεται στην υπολογισμένη ταχύτητα για τον υπολογισμό της νέας παραμέτρου. Συμπεραίνεται λοιπόν ότι η ποσότητα της ορμής επηρεάζεται από την μεταβολή του ρυθμού μάθησης [31].

Η ανάγκη που προκύπτει εδώ είναι η τροποποίηση του αλγορίθμου αυτού, έτσι ώστε να εφαρμόζονται διαφορετικές τιμές ορμής και ρυθμού μάθησης (αυτοπεποίθηση/*momentum* και άγχος/*anxiety* αντίστοιχα) για τον υπολογισμό των νέων τιμών των συναισθηματικών παραμέτρων του δικτύου. Ακόμα, επιχειρείται να τροποποιηθεί ο τρόπος εφαρμογής του ρυθμού μάθησης, έτσι ώστε να προσεγγίζει περισσότερο τις εξισώσεις που αναφέρθηκαν στο Κεφάλαιο 3. Με τον τρόπο αυτό αποφεύγεται η κλιμάκωση της προηγούμενης αλλαγής της παραμέτρου βάσει του γινομένου του ρυθμού μάθησης και της ορμής. Επομένως, οι εξισώσεις ανανέωσης των μεταβλητών προσαρμόζονται ως εξής:

$$u = \rho * u + lr * g$$

και

$$p = p - u$$

Εδώ εναλλάσσονται οι ποσότητες lr, ρ με τις m, k , δηλαδή του άγχους (anxiety) και της αυτοπεποίθησης (confidence) σε περίπτωση που οι παράμετροι που ανανεώνονται είναι οι συναισθηματικές. Η ανανέωση της αυτοπεποίθησης και του άγχους πραγματοποιείται σε κάθε επανάληψη βάσει των εξισώσεων 4.3.2.1.1 και 4.3.2.1.3 του Κεφαλαίου 4.

Οι τιμές που δόθηκαν σε αυτές τις σταθερές για το ORL dataset είναι αυτές που αναφέρονται από τους Dai et al., δηλαδή $lr = 0.2$ και $\rho = 0.41$. Στην περίπτωση του KDEF dataset οι αντίστοιχες τιμές είναι $lr = 0.01$ και $\rho = 0.1$. Παρατηρείται ότι οι τιμές των σταθερών αυτών στη δεύτερη περίπτωση συνόλου δεδομένων είναι σημαντικά μικρότερες. Μερικοί λόγοι στους οποίους μπορεί να αποδοθεί η σύγκλιση του δικτύου για μικρότερες τιμές παραμέτρων είναι η online μάθηση και επομένως η μεγάλη διακύμανση από επανάληψη σε επανάληψη εκπαίδευσης, καθώς και τα περισσότερα χαρακτηριστικά εισόδου που λαμβάνει το δίκτυο λόγω της φύσης του συνόλου δεδομένων KDEF.

5.3 Αποτελέσματα

Εδώ παρουσιάζεται η απόδοση του μοντέλου στα δύο επιλεγμένα σύνολα δεδομένων σε συνδυασμό με ανάλυση ευαισθησίας της απόδοσης του δικτύου με τροποποιήσεις στις εποχές και το μέγεθος του συνόλου αξιολόγησης (test set).

5.3.1 Εφαρμογή στο ORL σύνολο δεδομένων

Το μοντέλο στην ερευνητική εργασία των Dai et al. [9] εφαρμόζεται στο ORL σύνολο δεδομένων. Σε πρώτο στάδιο το μοντέλο που υλοποιήθηκε με PyTorch εφαρμόζεται στα ίδια δεδομένα. Τα αποτελέσματα αυτής της εφαρμογής παρουσιάζονται παρακάτω.

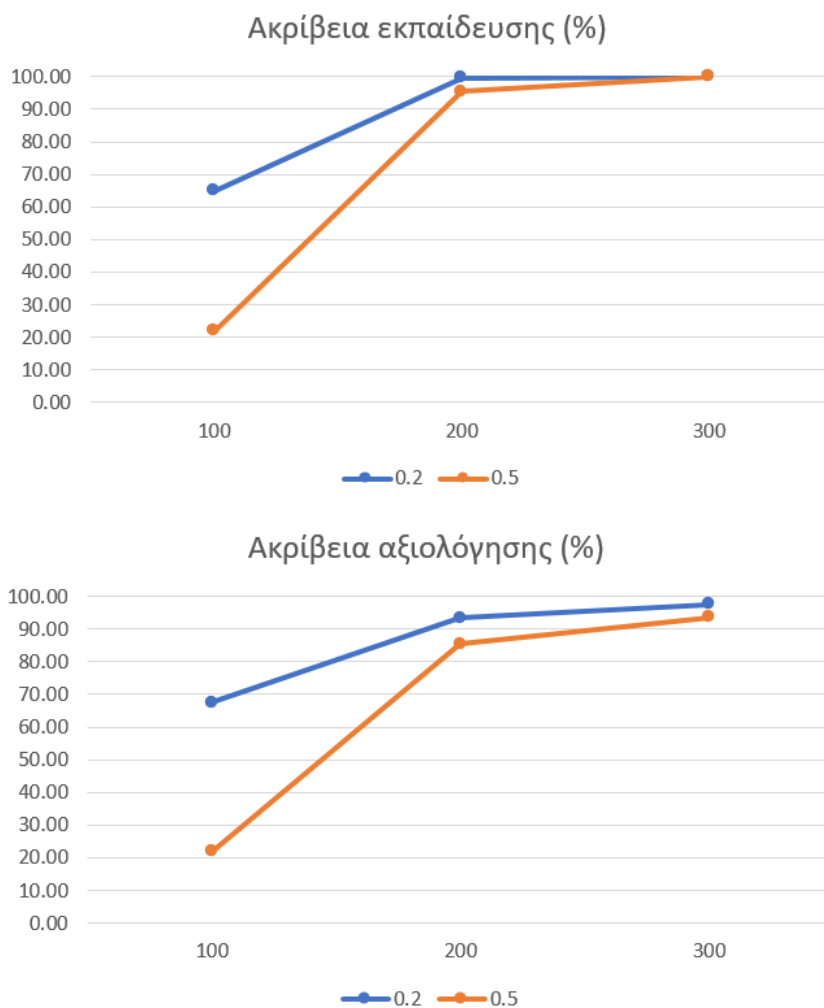
epochs / test size	0.2	0.5
100	64.93%	24.50%
200	99.64%	95.50%
300	100.00%	100.00%

Πίνακας 5.1: Ποσοστά ορθότητας (accuracy) εκπαίδευσης για διαφορετικές τιμές εποχών και μεγέθους test δεδομένων

epochs / test size	0.2	0.5
100	67.50%	22.00%
200	93.33%	85.50%
300	97.50%	93.50%

Πίνακας 5.2: Ποσοστά ορθότητας (accuracy) αξιολόγησης για διαφορετικές τιμές εποχών και μεγέθους test δεδομένων

Τα παραπάνω δεδομένα απεικονίζονται γραφικά στην Εικόνα 5.3. Όπως είναι φανερό η καλύτερη ορθότητα (accuracy) αξιολόγησης επιτυγχάνεται στις 300 εποχές όπου το μέγεθος των δεδομένων αξιολόγησης καταλαμβάνει το 20% των συνολικών δεδομένων.



Εικόνα 5.3: Διαγράμματα ορθότητας (accuracy) εκπαίδευσης και αξιολόγησης στο ORL dataset

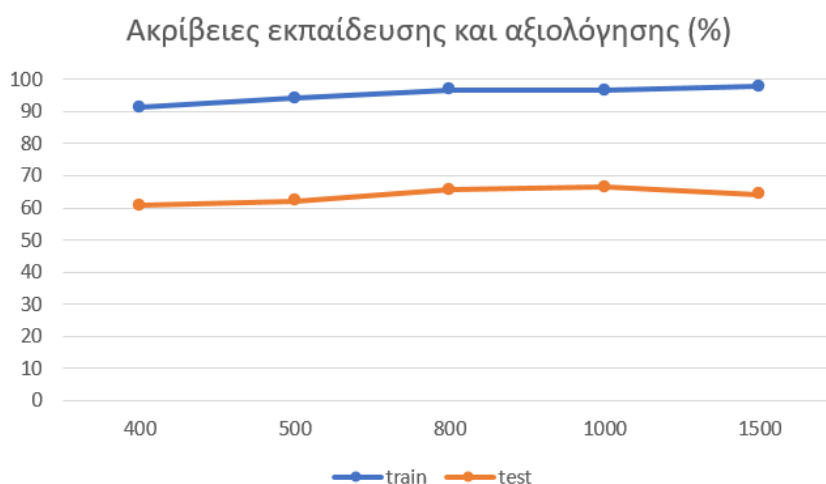
5.3.2 Εφαρμογή στο KDEF σύνολο δεδομένων

Τα αποτελέσματα ταξινόμησης του μοντέλου στο KDEF σύνολο δεδομένων δεν είναι τόσο ικανοποιητικά. Πιο συγκεκριμένα, παρά τις δοκιμές διαφορετικών τιμών παραμετροποίησης και εποχών εκπαίδευσης η ορθότητα (accuracy) αξιολόγησης στο test σύνολο δεδομένων δεν ξεπέρασε το 66-67%. Η ορθότητα αυτή προκύπτει στις 1000 εποχές εκπαίδευσης για τις τιμές παραμέτρων του υβριδικού δικτύου που αναφέρθηκαν και παραπάνω. Για 1500 εποχές το δίκτυο φαίνεται να μαθαίνει μεγαλύτερο ποσοστό θορύβου από εκείνον που περιέχεται στα δεδομένα, με αποτέλεσμα να χάνει την ικανότητα του να γενικεύει και να αποδίδει στα δεδομένα αξιολόγησης (φαινόμενο υπερεκπαίδευσης, *overfitting*). Αυτό αποτυπώνεται στην Εικόνα 5.4 όπου για 1500 εποχές εκπαίδευσης η ορθότητα εκπαίδευσης (train accuracy) αυξάνεται, ενώ η ορθότητα αξιολόγησης (test accuracy) μειώνεται σε σχέση με τις αντίστοιχες τιμές των 1200 εποχών.

Η απόδοση αυτή του δικτύου στο νέο σύνολο δεδομένων μπορεί να ερμηνευτεί ως κάποιος βαθμός από την προεπεξεργασία των δεδομένων, αλλά και τη δομή του δικτύου. Όσον αφορά την προεπεξεργασία των δεδομένων, η μείωση του μεγέθους των εικόνων κατά μεγάλο παράγοντα σμίχρυνσης οδηγεί στην απώλεια πολλών χαρακτηριστικών και προτύπων από την πληροφορία που εμπεριέχει η κάθε εικόνα. Επομένως, λόγω του μεγάλου

epochs	train	test
400	91.38%	60.82%
500	94.10%	62.24%
800	96.80%	65.71%
1000	96.51%	66.53%
1500	97.76%	64.29%

Πίνακας 5.3: Ποσοστά ορθότητας για διαφορετικές τιμές εποχών



Εικόνα 5.4: Ορθότητα (accuracy) εκπαίδευσης και αξιολόγησης ως προς τις εποχές εκπαίδευσης για το KDEP σύνολο δεδομένων

μεγέθους των διαθέσιμων εικόνων σε αυτό το σύνολο δεδομένων και του μεγάλου παράγοντα συρρίκνωσης τους, χάνεται αναπόφευκτα μεγάλο ποσοστό της πληροφορίας που το μοντέλο μπορεί να "μάθει" και να γενικεύσει σε άγνωστες εικόνες. Μία πιθανή λύση σε αυτό το πρόβλημα θα μπορούσε να αποτελέσει η αναγνώριση προτύπων και χαρακτηριστικών που φέρουν πληροφορία για τις συναισθηματικές εκφράσεις, οι οποίες να τροφοδοτηθούν στο δίκτυο αντί ολόκληρων των εικόνων. Η δοκιμή αυτή αφήνεται στον αναγνώστη ως περαιτέρω επέκταση.

Επιπρόσθετα, η δομή του υβριδικού δικτύου δεν συμβάλει στην ικανότητα του να γενικεύει στα δεδομένα με τα οποία τροφοδοτείται. Ο συνδυασμός της σιγμοειδούς συνάρτησης ενεργοποίησης με το μέσο τετραγωνικό σφάλμα ως συνάρτησης κόστους οδηγεί στο φαινόμενο του κορεσμού του δικτύου (**saturation**). Αυτός ο συνδυασμός καθιστά την συνάρτηση απώλειας μη κυρτή, ενώ παρατηρείται κορεσμός στην ανανέωση των παραμέτρων (**gradient saturation**). Το πρόβλημα εδώ είναι ότι μία μία πιθανά μεγάλη απόκλιση ανάμεσα στην πρόβλεψη του δικτύου και την τιμή στόχο μπορεί να οδηγήσει σε μία σχεδόν ανύπαρκτη ανανέωση των παραμέτρων του δικτύου. Το αποτέλεσμα είναι ότι το δίκτυο σταματάει να μαθαίνει και εγκλωβίζεται σε κάποιο από τα αρκετά τοπικά ελάχιστα που εμφανίζει μία μη κυρτή συνάρτηση κόστους. Έχει ακόμα αποδειχθεί ότι υπάρχει συσχέτιση ανάμεσα στον έντονο κορεσμό και την υπερεκπαίδευση και, συνεπώς, με τον έλεγχο του βαθμού κορεσμού του δικτύου ενισχύεται και η ικανότητα γενίκευσης του [2].

Κεφάλαιο 6

Συμπεράσματα

Στο παρόν κεφάλαιο αναπτύσσονται τα συμπεράσματα που προέκυψαν από την υλοποίηση του συναισθηματικά επαυξημένου υβριδικού δικτύου EmSOM, καθώς και πιθανές επεκτάσεις που θα μπορούσαν να αξιοποιηθούν για περαιτέρω μελέτη και βελτιστοποίηση ενός τέτοιου νευρωνικού δικτύου.

6.1 Συμπεράσματα υλοποίησης

Με την υλοποίηση του νευρωνικού δικτύου επιτεύχθηκε σημαντική απόδοση στο σύνολο δεδομένων ORL που στοχεύει στην αναγνώριση προσώπων. Η επιτυχία του νευρωνικού αυτού δικτύου σε δεδομένα εικόνων είναι ιδιαίτερα σημαντική αφού παρακάμπτονται τα συνελκτικά νευρωνικά δίκτυα που, λόγω πολλών ετερόκλητων επιπέδων (συνέλιξης, δειγματοληπτικά, ολικά συνδεδεμένα), κοστίζουν συνήθως περισσότερο υπολογιστικά και σε αριθμό παραμέτρων δικτύου (*coefficients*). Επομένως, το εν λόγω υβριδικό δίκτυο καταφέρνει με μικρότερο αριθμό παραμέτρων την σχεδόν τέλεια αναγνώριση προσώπων. Χρειάζεται ωστόσο η απεικόνιση των τιμών των *pixels* κάθε εικόνας στο εύρος 0 έως 0.1, έτσι ώστε να μειωθεί το μέτρο των αποκλίσεων του μοντέλου στα πλαίσια της ευαισθησίας του δικτύου στον κορεσμό ανανέωσης των παραμέτρων του.

Βάσει της επιτυχίας εφαρμογής του σε σύνολο δεδομένων εικόνων, το ίδιο δίκτυο επιχειρήθηκε να αξιολογηθεί σε δεδομένα εικόνων για αναγνώριση συναισθήματος βάσει εκφράσεων προσώπου. Στο επιλεγμένο σύνολο δεδομένων, ωστόσο, λόγω του μεγάλου μεγέθους εικόνων χρειάστηκε να επαναπροσδιοριστεί το μέγεθός τους σε μικρότερες διαστάσεις, έτσι ώστε να περιοριστεί ο αριθμός των βαρών ανάμεσα στο επίπεδο εισόδου και το κρυμμένο επίπεδο. Πέρα από την εξοικονόμηση υπολογιστικού κόστους, η μείωση των παραμέτρων του δικτύου μειώνει και την διακύμανσή του (*variance*), η οποία συμβάλλει στην υπερεκπαίδευσή του στα διαθέσιμα δεδομένα. Επειδή, λοιπόν, το επιλεγμένο μοντέλο είναι ευαίσθητο στην υπερπροσαρμογή στα δεδομένα λόγω κορεσμού, η διακύμανση του κρίθηκε σημαντικό να παραμείνει σε χαμηλά επίπεδα περιορίζοντας τον αριθμό των παραμέτρων. Παρά την απώλεια χαρακτηριστικών λόγω της συρρίκνωσης των εικόνων, η επίδοση του δικτύου είναι αξιόλογη και θα μπορούσε να βελτιωθεί περαιτέρω με εφαρμογή τεχνικών αναγνώρισης προτύπων (*pattern recognition*). Οι τεχνικές αυτές εστιάζουν στα χαρακτηριστικά κάθε δεδομένου εισόδου που φέρουν ουσιαστική πληροφορία ως προς τη ζητούμενη διεργασία (π.χ. *face landmarking*). Με τον τρόπο αυτό είναι δυνατό να επιτευχθεί μεγαλύτερη ορθότητα ταξινόμησης σε ένα από τα επτά συναισθήματα, αφού η εστίαση στα χαρακτηριστικά που φέρουν αξιοποιήσιμη από το δίκτυο πληροφορία, χωρίς την αύξηση του πλήθους των παραμέτρων του, παρέχει περισσότερα ερείσματα στην ανάπτυξη της ικανότητας του μοντέλου να γενικεύει.

6.2 Μελλοντικές επεκτάσεις

Ένα βασικό δομικό στοιχείο των περισσότερων δικτύων που εφαρμόζουν την συναισθηματική οπισθοδιάδοση (**emotional backpropagation**) είναι ο συνδυασμός της σιγμοειδούς συνάρτησης ενεργοποίησης στο επίπεδο εξόδου με το μέσο τετραγωνικό σφάλμα (**mean square error, MSE**) για τον υπολογισμό της απώλειας στις προβλέψεις του δικτύου. Βάσει αυτών των δικτύων δημιουργήθηκε και το εξεταζόμενο σε αυτή τη διπλωματική εργασία μοντέλο, το οποίο παρουσιάζει τα ίδια δομικά χαρακτηριστικά. Όπως αναλύθηκε και παραπάνω, οι δομικές αυτές επιλογές του δικτύου καθιστούν την συνάρτηση απώλειας μη κυρτή, με αποτέλεσμα να υπάρχουν πολλαπλά τοπικά ελάχιστα, στα οποία το μοντέλο παγιδεύεται. Αυτό το γεγονός κατέχει μεγάλο αντίκτυπο στην απόδοση και την ικανότητα γενίκευσης του δικτύου.

Σύμφωνα με τα παραπάνω, μία πρόταση βελτιστοποίησης είναι η εφαρμογή της συναισθηματικής οπισθοδιάδοσης (**emotional backpropagation**) αλλά και της αναπαράστασης της συναισθηματικής εισόδου με χάρτες αυτο-οργάνωσης σε διαφορετικής δομής δίκτυα. Ένα εναλλακτικό νευρωνικό δίκτυο θα μπορούσε να επιλέξει διαφορετική συνάρτηση απώλειας (**loss function**), όπως τη διασταυρωμένη εντροπία (**cross-entropy loss**). Οι λογαριθμικοί όροι σε αυτή τη συνάρτηση κόστους αποτρέπουν την εκδήλωση του φαινομένου του κορεσμού, αφού ακόμα και αν το μοντέλο δεν έχει καλή απόδοση, η συνάρτηση αυτή ωθεί το δίκτυο να συγκλίνει στο ολικό ελάχιστο.

Συνεπώς, με την αλλαγή της συνάρτησης κόστους ή της συνάρτησης ενεργοποίησης του επιπέδου εξόδου του δικτύου είναι αναμενόμενος ο κατευνασμός της ευασθησίας κορεσμού του μοντέλου και η καθιέρωση ενός δικτύου που αξιοποιεί τα χαρακτηριστικά των δεδομένων εισόδου που φέρουν πληροφορία προς ουσιαστική μάθηση. Σε αυτή την περίπτωση η προεπεξεργασία των δεδομένων θα είναι διαφορετική, αφού δεν θα είναι απαραίτητη η κανονικοποίηση των τιμών των **pixels** σε μικρότερο εύρος προς περιορισμό του μέτρου απόκλισης ανάμεσα στην επιθυμητή τιμή και την τιμή που έχει εκτιμηθεί από το μοντέλο. Τέλος, αναμενόμενη είναι σε αυτή την περίπτωση και η αύξηση της ορθότητας ταξινόμησης του μοντέλου ανεξαρτήτου συνόλου δεδομένων.

Βιβλιογραφία

- [1] D. Tranel-A.R. Damasio A. Bechara H. Damasio. “Deciding advantageously before knowing the advantageous strategy”. In: *Science* 275 (5304 1997), pp. 1293–1295. DOI: 10.1126/science.275.5304.1293.
- [2] A. Engelbrecht A. Rakitianskaia. “Measuring Saturation in Neural Networks”. In: *IEEE Symposium Series on Computational Intelligence*. IEEE, 2015.
- [3] Dr. C. George Boeree. *The Emotional Nervous System*. URL: <https://webpace.ship.edu/cgboer/limbicsystem.html>.
- [4] J. Moréan C. Balkenius. “Emotional Learning: A Computational Model of the Amygdala”. In: *Cybernetics and Systems: An International Journal* 32 (6 2001), pp. 611–636. DOI: 10.1080/01969720118947.
- [5] P. Gaussier C. Hasson and S. Boucenna. “Emotions as a dynamical system: the interplay between the meta-control and communication function of emotions”. In: *Paladyn* 2.3 (2011), pp. 111–125. DOI: 10.2478/s13230-012-0005-4.
- [6] P. Gaussier C. Hasson and S. Boucenna. “Robots have emotions”. In: *Adaptive Behavior* 18.6 (2010), pp. 453–469. DOI: 10.1177/1059712310388528.
- [7] D. Shahmirzadi C. Lucas and N. Sheikholeslami. “Introducing BELBIC: Brain emotional learning based intelligent controller”. In: *Intelligent Automation Soft Computing* 10.1 (2004), pp. 11–21. DOI: 10.1080/10798587.2004.10642862.
- [8] D. Shahmirzadi C. Lucas and N. Sheikholeslami. “Practical emotional neural networks”. In: *Neural Networks* 59 (Nov. 2014), pp. 61–72. DOI: 10.1016/j.neunet.2014.06.012.
- [9] Q. Dai and L. Guo. “Two novel hybrid Self-Organizing Map based emotional learning algorithms”. In: *Neural Computing and Applications* 31 (Oct. 2019), 2921–2938. DOI: 10.1007/s00521-017-3240-0.
- [10] A.R. Damasio. *Descartes’ Error: Emotion, Reason, and the Human Brain*. New York, NY: Gosset/Putnam Press, 1994.
- [11] M.-R. Akbarzadeh-T. E. Lotfi. “Adaptive brain emotional decayed learning for online prediction of geomagnetic activity indices”. In: *Neurocomputing* 126 (Feb. 2014), pp. 188–196. DOI: 10.1016/j.neucom.2013.02.040.
- [12] M.-R. Akbarzadeh-T. Ehsan Lotfi. “Brain emotional learning-based pattern recognizer”. In: *Cybernetics and Systems* 44.5 (2013), pp. 402–421. DOI: 10.1080/01969722.2013.789652.
- [13] M.-R. Akbarzadeh-T. Ehsan Lotfi. “Emotional brain-inspired adaptive fuzzy decayed learning for online prediction problems”. In: Hyderabad, India: IEEE, 2013, pp. 1–7. DOI: 10.1109/FUZZ-IEEE.2013.6622510.
- [14] Eklavya. *Kohonen Self-Organizing Maps*. URL: <https://towardsdatascience.com/kohonen-self-organizing-maps-a29040d688da>.

- [15] *The Karolinska Directed Emotional Faces - KDEF*. Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, 1998. ISBN: 91-630-7164-9. URL: <https://www.kdef.se>.
- [16] “Emotion-Augmented Machine Learning: Overview of an Emerging Domain”. In: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. San Antonio, TX, USA: IEEE, Oct. 2017, pp. 305–312. DOI: 10.1109/ACII.2017.8273617.
- [17] “Emotion-driven reinforcement learning”. In: *Annual Meeting of the Cognitive Science Society*. Vol. 30. 30. Wahington DC: Cognitive Science Society, 2008.
- [18] W. Lv L. Liu V. G. Duffy F. Guo F. Li. “Bibliometric Analysis of Affective Computing Researches during 1999–2018”. In: *International Journal of Human–Computer Interaction* 36 (9), pp. 801–814. DOI: 10.1080/10447318.2019.1688985.
- [19] J. Healey and R. Picard. “Digital processing of affective signals”. In: *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98*. Vol. 6. Seattle, WA, USA: IEEE, 1998, pp. 3749–3752. DOI: 10.1109/ICASSP.1998.679699.
- [20] D.O. Hebb. *A Textbook of Psychology*. W.B. Saunders Co., 1966.
- [21] Timo Honkela. *Description of Kohonen’s Self-Organizing Map*. URL: <http://www.mlab.uiah.fi/~timo/som/thesis-som.html>.
- [22] Queensland Brain Institute. *The limbic system*. URL: <https://qbi.uq.edu.au/brain/brain-anatomy/limbic-system>.
- [23] J. . R. Jang. “ANFIS: Adaptive-network-based fuzzy inference system”. In: *Transactions on Systems, Man, and Cybernetics* 23.3 (1993), pp. 665–685. DOI: 10.1109/21.256541.
- [24] A. Khashman. *A Modified Backpropagation Learning Algorithm With Added Emotional Coefficients*. Nov. 2008. DOI: 10.1109/TNN.2008.2002913.
- [25] A. Khashman. *Application of an emotional neural network to facial recognition*. Nov. 2009. DOI: 10.1109/TNN.2008.2002913.
- [26] A. Khashman. *Modeling cognitive and emotional processes: A novel neural network architecture*. Dec. 2010. DOI: 10.1016/j.neunet.2010.07.004.
- [27] A. Khashman. “Modeling cognitive and emotional processes: A novel neural network architecture”. In: vol. 23. Elsevier, 2010, pp. 1155–1163. DOI: 10.1016/j.neunet.2010.07.004.
- [28] Achraf Khazri. *Self-Organizing Maps*. URL: <https://towardsdatascience.com/self-organizing-maps-1b7d2a84e065>.
- [29] Teuvo Kohonen. *The Self-Organizing Map*. Sept. 1990. DOI: 10.1109/5.58325.
- [30] J.E. LeDoux. *The emotional brain*. Simon and Schuster, 1996.
- [31] Ceshine Lee. *SGD implementation in PyTorch*. 2017. URL: <https://medium.com/the-artificial-impostor/sgd-implementation-in-pytorch-4115bcb9f02c>.
- [32] D. E. Rumelhart; J. L. McClelland. “Learning Internal Representations by Error Propagation”. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*. 1987, pp. 318–362.
- [33] “Modeling emotions and other motivations in synthetic agents”. In: *AAAI/I-AAI*. Providence, Rhode Island, 1997, pp. 10–15. DOI: 10.5555/1867406.1867408.

- [34] *ORL dataset (Our Database of Faces)*. AT&T Laboratories Cambridge, 2002. URL: <https://cam-orl.co.uk/facedatabase.html>.
- [35] Benjamin H. Ott. *A convergence criterion for Self-Organizing Maps*. 2012. URL: <https://homepage.cs.uri.edu/faculty/hamel/pubs/theses/ms-thesis-ben.pdf>.
- [36] F. Melo P. Sequeira and A. Paiva. “Emotion-based intrinsic motivation for reinforcement learning agents”. In: *Affective Computing and Intelligent Interaction*. Springer, 2011, pp. 326–336. DOI: 10.1007/978-3-642-24600-5_36.
- [37] F. Melo P. Sequeira and A. Paiva. “Learning by appraising: An emotion-based approach to intrinsic reward design”. In: *Adaptive Behavior* 22.5 (2014), pp. 330–349. DOI: 10.1177/1059712314543837.
- [38] R. W. Picard. *Affective Computing*. Tech. rep. 321. MIT Media Laboratory Perceptual Computing Section, 1995.
- [39] R. W. Picard. “Affective computing: challenges”. In: *International Journal of Human-Computer Studies* (2013). DOI: 10.1016/S1071-5819(03)00052-1.
- [40] R. W. Picard. “Affective Computing for HCI”. In: *Proceedings of HCI International (the 8th International Conference on Human-Computer Interaction)*. 1999, pp. 829–833.
- [41] R. W. Picard. *Affective Computing: From Laughter to IEEE*. 2010. DOI: 10.1109/T-AFFC.2010.10.
- [42] A. Kumar R. Arya J. Singh. In: *A survey of multidisciplinary domains contributing to affective computing* 40 (2021). DOI: 10.1016/j.cosrev.2021.100399.
- [43] S. Baron-Cohen R. Kaliouby R. Picard. In: *Affective computing and autism* 1093 (2006), pp. 228–248. DOI: 10.1196/annals.1382.016.
- [44] H. Roibu R. Popescu D. Ponescu and L. Popescu. “Smart Classroom - Affective Computing in Present-Day Classroom”. In: *2018 28th EAEEIE Annual Conference (EAEEIE)*. Hafnarfjordur, Iceland: IEEE, 2018, pp. 1–9. DOI: 10.1109/EAEEIE.2018.8534286.
- [45] P. Zahadat R. Thenius and T. Schmickl. “EMANN – a model of emotions in an artificial neural network”. In: *European Conference on Artificial Life*. Vol. 12. 11. Sept. 2013, pp. 830–837. DOI: 10.7551/978-0-262-31709-2-ch122.
- [46] W. Bender et al. R.W. Picard S. Papert. “Affective Learning — A Manifesto”. In: *BT Technology Journal* 22 (2004). DOI: 10.1023/B:BTTJ.0000047603.37042.33.
- [47] Scholarpedia. *Three postulates of Hebb’s theory*. URL: http://www.scholarpedia.org/article/Donald_Olding_Hebb.
- [48] D. Schuller and B. W. Schuller. *The Age of Artificial Emotional Intelligence*. Sept. 2018. DOI: 10.1109/MC.2018.3620963.
- [49] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An introduction*. The MIT Press, 2011, p. 334.
- [50] R. Karimizandi T. Babaie and C. Lucas. “Learning based brain emotional intelligence as a new aspect for development of an alarm system”. In: *Soft Computing – A Fusion of Foundations, Methodologies and Applications* 12.9 (2008), pp. 857–873. DOI: 10.1007/s00500-007-0258-8.

- [51] L. Jäncke T. Baumgartner M. Esslen. “From emotion perception to emotion experience: Emotions evoked by pictures and classical music”. In: *International Journal of Psychophysiology* 60 (1 Apr. 2006), pp. 34–43. DOI: 10.1016/j.ijpsycho.2005.04.007.
- [52] Q.-H. Vuong T. Ho H.-K. Nguyen. “Affective computing at the edge: A bibliometric analysis of the period 1995-2020”. Unpublished manuscript with DOI: 10.31219/osf.io/hu54t.
- [53] N. Cooch T. Stalnaker and G. Schoenbaum. “What the orbitofrontal cortex does not do”. In: *Nature Neuroscience* 18 (5 May 2015), pp. 620–626. DOI: 10.1038/nn.3982.
- [54] J. D. Velasquez. *Modeling emotion-based decision-making*. Tech. rep. 1998, pp. 164–169.
- [55] J. Wallis. “Cross-species studies of orbitofrontal cortex and value-based decision-making”. In: *Nature Neuroscience* 15 (2012), pp. 13–19. DOI: 10.1038/nn.2956.
- [56] P. Werbos. In: *Origins: Brain and Self-Organization*. Lawrence Erlbaum Associates, 1994. Chap. The brain as a neurocontroller: New hypotheses and experimental possibilities. Pp. 680–706.
- [57] D. H. Hubel T. N. Wiesel. “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex”. In: *The Journal of Physiology* 160 (1 Jan. 1962), pp. 106–154. DOI: 10.1113/jphysiol.1962.sp006837.
- [58] Wikipedia. *Prefrontal cortex*. URL: https://en.wikipedia.org/wiki/Prefrontal_cortex.
- [59] Wikipedia. *Sensory Cortex*. URL: https://en.wikipedia.org/wiki/Sensory_cortex.
- [60] X. Yuan Y. Yang Y. Wang and F. Yin. *Hybrid chaos optimization algorithm with artificial emotion*. 2012. DOI: 10.1016/j.neunet.2010.07.004.