

Υψηλού επιπέδου υλοποίηση των αλγορίθμων HTM
με έμφαση στη μαθηματική διατύπωση ως βάση για
περαιτέρω μελέτη

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Κωνσταντίνος Σαμαράς-Τσακίρης

2019-06-05

Περιεχόμενα

1	Εισαγωγή	3
2	Το μοντέλο HTM	4
2.1	Το πρόβλημα της πρόβλεψης ακολουθιών	4
2.2	Στοιχεία της Hierarchical Temporal Memory	5
2.2.1	Μοντέλο νευρώνα	6
2.2.2	Αραιές Διανεμημένες Αναπαραστάσεις (SDR)	7
2.2.3	Μοντέλο δικτύου	9
2.2.4	Encoders	9
2.2.5	Decoders	10
2.3	Βασικές λειτουργίες HTM	10
2.3.1	Sequence Memory	10
2.4	Αποτελέσματα - Προσομοιώσεις	11

Κεφάλαιο 1

Εισαγωγή

Κεφάλαιο 2

Το μοντέλο HTM

2.1 Το πρόβλημα της πρόβλεψης ακολουθιών

Τα σύνορα της τεχνητής νοημοσύνης εκτείνονται πέρα από το παραδοσιακό πρόβλημα της εκμάθησης ενός στατικού συνόλου δεδομένων με επιβλεπόμενες μεθόδους, όπως γίνεται σε πολλές σύγχρονες εφαρμογές [14]. Η τεχνητή νοημοσύνη καλείται σήμερα να αξιοποιήσει τις καταγιστικές ροές δεδομένων που παρέχει το (ανθρωπογενές και μη) περιβάλλον, σε πραγματικό χρόνο, και δίχως την πολυτέλεια της προσήμανσης που απαιτεί η επιβλεπόμενη μάθηση — εφαρμογές IoT που απαιτούν δυναμική αλληλεπίδραση με το περιβάλλον τους έρχονται στο μυαλό [1].

Επομένως, καλούμαστε να μοντελοποιήσουμε έναν κόσμο που αλλάζει [16]. Το μοντέλο οφείλει ή να είναι γενικότερο από όλες τις δυνατές μεταβολές του κόσμου ή να αλλάζει μαζί του. Η έννοια της ροής δεδομένων που αναφέρθηκε υπονοεί την έννοια του χρόνου, της αλληλουχίας, που επιτρέπει στο μοντέλο να αντλήσει πληροφορία από την αιτιώδη σχέση. Την ερμηνεία του κόσμου με τη βοήθεια της συνεπαγωγής και της αναγνώρισης απαιτήσεων και συνεπειών.

Ορίζεται έτσι το πρόβλημα της εκμάθησης και πρόβλεψης ακολουθιών. Δηλαδή, ο πράκτορας που παρακολουθεί μια αλληλουχία δεδομένων (γεγονότων) καλείται να προβλέψει τη συνέχεια και να δράσει έτσι, ώστε βέλτιστα να ανταμειφθεί. Η ανταμοιβή, το κίνητρο και γενικότερα ο στόχος έχουν τεθεί από εξωτερικό παράγοντα και βρίσκονται εκτός του πλαισίου του προβλήματος. Άμεση εφαρμογή της πρόβλεψης είναι και η αναγνώριση ανωμαλιών.

Η εκμάθηση ακολουθιών είναι κλασικό πρόβλημα, για το οποίο έχουν αναπτυχθεί κλασικές λύσεις. Κυριότερο και ευρύτερα διαδεδομένο είδος μοντέλων, ειδικά πριν τη σύγχρονη επάνοδο των νευρωνικών δικτύων, είναι τα Hidden Markov Models. Τα κλασικά νευρωνικά δίκτυα προσφέρουν στο πρόβλημα τα νευρωνικά δίκτυα χρονικής καθυστέρησης (TDNN). Η ουσιαστική συνεισφορά των κλασικών νευρωνικών δικτύων επιτυγχάνεται όμως με τα ανάδρομα δίκτυα (RNN), που γενίκευσαν τις παλαιότερες προσθιοδρομικές (feedforward) αρχιτεκτονικές ακριβώς για να μπορούν να αντιμετωπίσουν εγγενώς προβλήματα ακολουθιών. Ιδιαίτερης μνείας χρήζει η δομή «μακράς βραχυπρόθεσμης μνήμης» (LSTM). Από το 2015 έχει χρησιμοποιηθεί με επιτυχία σε ποικίλες εφαρμογές πολύ πιο σύνθετες από την εισαγωγή που παρουσιάζεται εδώ, όπως ως τεχνητή νοημοσύνη που παίζει το παιχνίδι StarCraft 2 [AlphaStar 22].

Σε αυτόν το χώρο λύσεων υπάρχουν παρόλα αυτά σημεία για βελτίωση. Οι περισσότερες εφαρμογές, συμπεριλαμβανομένου του AlphaStar, βασίζονται σε επιβλεπόμενη και ενισχυτική μάθηση, αφήνοντας ευρύ πεδίο εξερεύνησης για μη επιβλεπόμενα μοντέλα. Καθώς ο όγκος των ροών δεδομένων αυξάνεται εκθετικά [18], παρατηρείται αυξανόμενη ζήτηση για αλγορίθμους που προσαρμόζονται γρήγορα

στις εξελισσόμενες στατιστικές των δεδομένων τους (συνεχής μάθηση), έχοντας πρόσβαση μόνο σε μικρό χρονικό παράθυρο πληροφορίας. Στην προσπάθεια αυτή αναπτύσσονται ενεργά τεχνικές για βελτίωση της αποδοτικότητας δειγμάτων της μάθησης [όπως 19], ή για παράκαμψη του προβλήματος, όπως η μεταφορική μάθηση (transfer learning) [21].

Στην ανάλυση των διαφόρων τρόπων ορισμού της τεχνητής νοημοσύνης, ο Wang [6] αναγνωρίζει τη μέθοδο «από τη δομή» για την έρευνα σε συστήματα που εμπνέονται ή μιμούνται το βασικό παράδειγμα ευφυούς συστήματος που επιλύει διαρκώς το παραπάνω πρόβλημα: τον εγκέφαλο, και ειδικά το φλοιό. Παρατηρείται μια τάση στο χώρο των τεχνητών νευρωνικών δικτύων ανασκόπησης της επαφής τους με τη βιολογική πραγματικότητα τα τελευταία χρόνια, όπως στο [8]. Το πιο ηχηρό παράδειγμα είναι ο μηχανισμός κάψουλας που πρότειναν ο Hinton και συνεργάτες [17], [20].

Σε αυτό λοιπόν το πλαίσιο, η μελέτη της HTM κρίνεται ιδιαίτερα καίρια.

2.2 Στοιχεία της Hierarchical Temporal Memory

Παρακάτω θα αναφερθούμε πάλι σε στοιχεία νευροεπιστήμης. Καθώς ο σκοπός της περιγραφής είναι η κατανόηση των αλγορίθμων HTM, πρέπει να γίνει αποδεκτή μια παρέκκλιση από την αυστηρότητα, χάριν απλότητας. Σε ό,τι βιολογικό στοιχείο αναφερθεί, ο αναγνώστης ας έχει υπόψιν ότι η πραγματικότητα είναι πάντα πιο πολύπλοκη και ότι αυτά που γράφονται εδώ δεν είναι ακριβώς ορθά. Είναι, όμως, *χρήσιμες προσεγγίσεις*.

Η κεντρική θέση στην οποία βασίζεται η Hierarchical Temporal Memory τμηματοποιεί τον εγκεφαλικό φλοιό σε ένα ψηφιδωτό βασικών μονάδων επεξεργασίας, των **”φλοιικών στηλών” (cortical columns)**, που έχουν την ίδια δομή και εκτελούν τους **ιδίους αλγορίθμους, αλλά σε διαφορετικά δεδομένα**. Η ιδέα αυτή έχει μακρά ιστορία στη νευροεπιστήμη, με μια συγκεντρωτική επισκόπηση από Defelipe, Markram κ.ά [7] να την οριοθετεί και να παρουσιάζει την ευρεία χρήση και κακομεταχείριση του όρου. Για παράδειγμα, η εισαγωγική πρόταση περί **”ψηφιδωτού”** παραπάνω πρέπει να αντιμετωπιστεί μόνο ως πρώτη προσέγγιση, γιατί οι στήλες δε φέρονται να έχουν γεωμετρικά σαφή σύνορα μεταξύ τους, αλλά διάχυτες ζώνες μετάβασης. Πρεσβευτής και βασικός αποκρυσταλλωτής της ιδέας είναι ο Mountcastle [3], με την ιδέα να χαίρει τόσο αποδοχής [11], όσο και κριτικής ως προς τη χρησιμότητά της [4]. Παρόλα αυτά, εδώ θα την υιοθετήσουμε. Έτσι, ο εγκεφαλικός φλοιός αποδομείται σε πλειάδα μονάδων επεξεργασίας κοινής αρχής, και οι αλγόριθμοι που θα παρουσιαστούν περιγράφουν τη λειτουργία της μονάδας.

Η φλοιική στήλη είναι λοιπόν ένας πληθυσμός νευρώνων με κοινή συνδεσμολογία: λαμβάνουν το σήμα εισόδου από κοινές πηγές και στέλνουν το σήμα εξόδου σε κοινούς παραλήπτες. Συχνά, τέτοιοι παραλήπτες είναι άλλες φλοιικές στήλες. Συντάσσονται έτσι επεξεργαστικές **ιεραρχίες**, με τα πρώτα στάδια της ιεραρχίας να δημιουργούν απλούστερα μοντέλα για τον κόσμο από τα μετέπειτα (κυρίως γιατί τα μετέπειτα συγκεντρώνουν περισσότερη πληροφορία).

Οι νευρώνες στους οποίους αναφερόμαστε παραπάνω, οι πυραμιδοειδείς νευρώνες, βρίσκονται κάθε στιγμή σε 1 από 2 καταστάσεις: ενεργοί ή ανενεργοί. Η ενεργοποίησή τους (**”δυναμικό δράσης”**) ερεθίζει άλλους νευρώνες με τους οποίους συνδέονται και μπορεί να τους οδηγήσει σε ενεργοποίηση. Μπορούμε λοιπόν να περιγράψουμε την κατάσταση του φλοιού κάθε στιγμή ως το σύνολο των νευρώνων που είναι ενεργό. Προκύπτει ότι το σύνολο αυτό είναι πολύ μικρό ποσοστό του συνολικού νευρικού πληθυσμού, δηλαδή η ενεργοποίηση είναι **αραιή**, περίπου 2%.

Σημειώνεται ότι οι πυραμιδοειδείς νευρώνες είναι η μειοψηφία των νευρώνων στο φλοιό. Το βασικό κοινό τους χαρακτηριστικό είναι ότι στέλνουν τους άξονές τους μέσω της λευκής ύλης σε μακρινούς προορισμούς, για το οποίο και ονομάζονται κύριοι νευρώνες. Περισσότεροι είναι οι διάμεσοι νευρώνες (interneurons), των οποίων οι άξονες παραμένουν σε μικρή εμβέλεια, και θεωρείται ότι συμμετέχουν σε τοπικά κυκλώματα που εν τέλει ρυθμίζουν την ενεργοποίηση των κυρίων νευρώνων [5].

Το μοτίβο ενεργοποίησης ενός πληθυσμού νευρώνων της ίδιας στήλης αποτελεί στο πλαίσιο της θεωρίας HTM τη δομή δεδομένων του εγκεφάλου και ονομάζεται **αραιή διανεμημένη αναπαράσταση (SDR)**. Κάθε αίσθηση στέλνει SDR στις στήλες που την επεξεργάζονται· κάθε στήλη στέλνει SDR στους μύες ή σε άλλες στήλες ως έξοδο. Η είσοδος σε ένα μοντέλο HTM πρέπει επομένως να μεταφράζει τη φυσική ποσότητα που θέλουμε να επεξεργαστούμε σε SDR, με τη διαδικασία να ονομάζεται **κωδικοποίηση**. Αντίστοιχα, η ερμηνεία ενός SDR εξόδου ονομάζεται **αποκωδικοποίηση**.

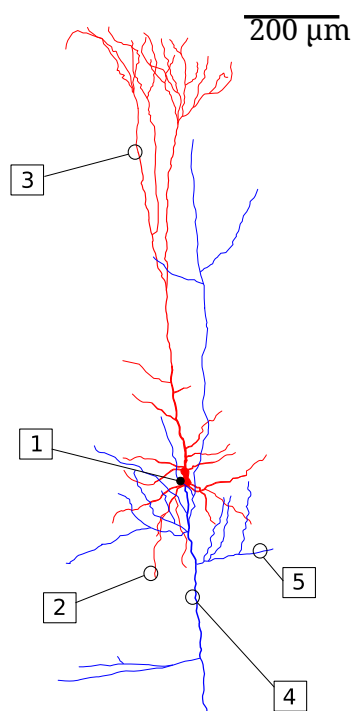
Από τους κοινούς αλγορίθμους που υλοποιεί κάθε φλοιική στήλη, η HTM αυτή τη στιγμή περιγράφει 2: τη **χωρική συγκέντρωση** και τη **χρονική μνήμη**. Η χρονική συγκέντρωση είναι επίσης διαδικασία που υποτίθεται, αλλά δεν έχει περιγραφεί επαρκώς και αποτελεί βασικό σημείο για περαιτέρω μελέτη.

2.2.1 Μοντέλο νευρώνα

Στον ανθρώπινο εγκέφαλο υπάρχουν πολλά είδη νευρώνων, που διαφέρουν στη μορφολογία, στις ηλεκτρικές και χημικές τους ιδιότητες. Η μόνη κατηγορία νευρώνα στην οποία βασίζεται το νευρικό μοντέλο της HTM είναι οι πυραμιδοειδείς νευρώνες 2.1 (το αποτέλεσμα της συμπεριφοράς άλλων νευρώνων συμπεριλαμβάνεται έμμεσα στη λογική των αλγορίθμων). Ο πυραμιδοειδής νευρώνας είναι ένα σύνθετο στοιχείο, με πολλούς χώρους και διανεμημένες λειτουργίες, που υλοποιούν τις λογικές πράξεις της άθροισης, του πολλαπλασιασμού, της χωρικής και χρονικής ολοκλήρωσης ταυτόχρονα και παράλληλα σε διαφορετικές ομάδες εισόδων. Η κεντρική δομή του νευρώνα, το σώμα, δέχεται ερεθίσματα από άλλους νευρώνες. Αρκετά τέτοια ερεθίσματα σε σύντομο χρονικό διάστημα είναι ικανά να ενεργοποιήσουν το νευρώνα. Οι πηγές των σημάτων που ο νευρώνας λαμβάνει στο σώμα του ονομάζονται ευαίσθητο πεδίο (receptive field) του νευρώνα. Από το σώμα φυτρώνουν οι εγγύς δενδρίτες. Ο ερεθισμός τους δεν είναι ικανός συνήθως να ενεργοποιήσει το νευρώνα, αρκεί όμως για να τον θέσει σε κατάσταση «επιφυλακής» (αποπολωμένος/προβλεπτικός), διευκολύνοντας τη μετέπειτα ενεργοποίησή του από σωματικά ερεθίσματα. Ομοίως και για τον ερεθισμό απομακρυσμένων ή κορυφαίων δενδριτών κατά μήκος του άξονα.

Στο πλαίσιο της HTM, νευρώνα ονομάζουμε μια δομή που δέχεται ερεθίσματα στους δενδρίτες της και τροποποιεί την τρισταθή της κατάσταση με βάση αυτά. Η κατάσταση μπορεί να είναι *ανενεργή*, *αποπολωμένη* (αλλιώς *προβλεπτική*) ή *ενεργή*. Όπως φαίνεται στο σχήμα 2.1β', ο νευρώνας HTM έχει 3 πύλες εισόδου. Ερεθίσματα στην προσθιοδρομική (feedforward) είσοδο προσμετρώνται και, αν ξεπεράσουν ένα κατώφλι, ο νευρώνας μεταβαίνει στην ενεργή κατάσταση. Ερεθίσματα στην είσοδο συμφραζομένων (context) είναι ικανά να προκαλέσουν αποπόλωση σε ανενεργό νευρώνα, αλλά όχι να τον ενεργοποιήσουν. Ομοίως και στην αναδρομική (feedback) είσοδο. Καλύπτονται έτσι μερικές συμπεριφορές του βιολογικού πυραμιδοειδούς νευρώνα. Σε αντιπαράβολή, ο νευρώνας ενός κλασικού τεχνητού νευρωνικού δικτύου δεν έχει διακριτά λειτουργικά τμήματα και απλώς αθροίζει όλες τις εισόδους του.

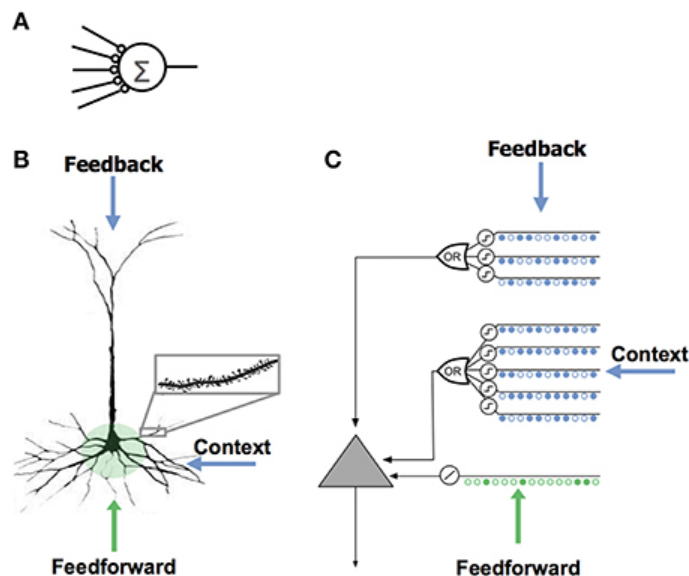
Η δομή που μεταφέρει σήματα μεταξύ νευρώνων ονομάζεται *σύναψη*. Συνάψεις



(α') Τυπική μορφή πυραμιδοειδούς νευρώνα:

1. σώμα
2. βασικός δένδριτης
3. κορυφαίος δένδριτης
4. άξονας
5. παράπλευρος άξονας

[πηγή 23]



(β')

Α'. Μοντέλο τυπικού νευρώνα σε τεχνητό νευρωνικό δίκτυο

Β'. Αναπαράσταση πυραμιδοειδούς νευρώνα

Γ'. Μοντέλο νευρώνα HTM

[πηγή 12]

υπάρχουν σε όλες τις ζώνες ερεθισμού που αναφέρθηκαν: στο σώμα, στους εγγύς και κορυφαίους δένδριτες. Οι συνάψεις είναι επίσης ο στόχος των μηχανισμών μάθησης: με διαδικασίες που θα αναλυθούν αργότερα, αλλά υπό την ομπρέλα του "Hebbian learning", το μέγεθος των συνάψεων αυξομειώνεται. Στο πλαίσιο της HTM όμως αυτό **δε συνεπάγεται συναπτικό βάρος**: οι συνάψεις είναι συνδεδεμένες ή αποσυνδεδεμένες. Μια σύναψη με πολύ μικρό μέγεθος θεωρείται αποσυνδεδεμένη και δε μετράει καθόλου ερεθίσματα. Καθώς μεγαλώνει, εφόσον περάσει ένα κατώφλι γίνεται συνδεδεμένη και μετράει ερεθίσματα.

Ένας πυραμιδοειδής νευρώνας μπορεί να έχει χιλιάδες συνάψεις. Ελάχιστες από αυτές βρίσκονται κοντά στο σώμα και, όπως προαναφέρθηκε, μπορούν να τον ενεργοποιήσουν. Οι υπόλοιπες βρίσκονται στους δένδριτες και, μεμονωμένα, έχουν πολύ μικρή επίδραση στο σώμα. Οι δένδριτες όμως προκύπτει ότι παίζουν ρόλο σύνθετων επεξεργαστικών στοιχείων. Αν ενεργοποιηθούν *περισσότερες από 8-20 συνάψεις* σε μικρό χρόνο στον ίδιο δένδριτη, τότε ο δένδριτης ενεργοποιείται και δημιουργεί ένα δυναμικό που προκαλεί την αποπόλωση όλου του νευρώνα [12]. Έτσι, κάθε δένδριτης λειτουργεί ως «**ανιχνευτής συμπτώσεων**» (coincidence detector).

Ας εξερευνήσουμε γιατί.

2.2.2 Αραιές Διανεμημένες Αναπαραστάσεις (SDR)

Με βάση την παρατήρηση της αραιής ενεργοποίησης των νευρώνων στον εγκέφαλο σχεδιάζουμε ένα μεγάλο, αραιό, δυαδικό διάνυσμα που αναπαριστά την κατάσταση ενεργοποίησης κάθε νευρώνα της περιοχής που μελετούμε. Ουσιαστικά, αυτά τα αραιά διανύσματα αποτελούν τη «δομή δεδομένων» [2], [13] του εγκεφάλου.

Αναφέρθηκε στην 2.2.1 ότι οι δενδρίτες λειτουργούν ως ανιχνευτές συγκεκριμένων μοτίβων ενεργοποίησης. Πώς όμως αρκούν 8-20 συνάψεις για να διακρίνουν συγκεκριμένα μοτίβα μέσα σε μεγάλους νευρικούς πληθυσμούς; Το κλειδί είναι η αραιή ενεργοποίηση.

Έστω ένας πληθυσμός 200K νευρώνων, όπου ενεργοί είναι το 1%, και ένας δενδρίτης που απαιτεί 10 ερεθίσματα για να ενεργοποιηθεί. Αν τυγχάνει να έχει συνδεδεμένες συνάψεις σε 10 από τους 2000 ενεργούς νευρώνες, ενεργοποιείται. Το «μοτίβο» εν προκειμένω είναι το σύνολο των συγκεκριμένων 2000 νευρώνων που ενεργοποιήθηκαν. Προφανώς, αφού έχει συνάψεις μόνο με 10/2000 ενεργούς νευρώνες, ο δενδρίτης θα μπορούσε να ενεργοποιηθεί κατά λάθος και σε πολλά διαφορετικά μοτίβα, που τυχαίνει να μοιράζονται τους ίδιους 10 νευρώνες με το αρχικό. Πόση είναι η πιθανότητα ενός τέτοιου σφάλματος; 9.8×10^{-21} .

Παρακάτω θα μελετήσουμε τις ιδιότητες των SDR για να διαπιστωθεί πώς προκύπτει αυτό το αποτέλεσμα.

Χωρητικότητα SDR

Έστω N-bit SDR $s = 0, 1^N$. Ο αριθμός των μονάδων στο s $w = count(s)$ ονομάζεται πληθάριθμος. Η χωρητικότητα SDR με μέγεθος N και πληθάριθμο w είναι το πλήθος των διαφορετικών SDR με αυτή τη μορφή, δηλαδή οι συνδυασμοί των N ανά w:

$$\text{χωρητικότητα}(N, w) = \binom{N}{w} = \frac{N!}{w!(N-w)!}$$

Συνδυασμοί SDR

Έστω 2 SDR A, B μήκους N. Ορίζουμε:

Union	$A B$
Overlap	$A\&B$
Overlap score	$\ A\&B\ $
Overlap set(θ)	$\{K \text{ όπου } \ A\&K\ > \theta\}$
Ταιριάζουν(θ)	A, B ταιριάζουν, αν ανήκουν στο ίδιο overlap set(θ)

Ας μελετήσουμε τον ορισμό ότι 2 SDR A με πληθάριθμο w και B *ταιριάζουν*. Έστω $B := A + 30\%$ θόρυβος τυχαίας αναστροφής bit. Τότε το προσδοκώμενο overlap score A, B θα είναι $70\%w$. Αν $\theta := 70\%w$, τα A, B προσδοκείται να ταιριάζουν.

Το παράδειγμα αυτό μπορεί να εστιαστεί για την περίπτωση της ενότητας 2.2.2. Έστω A το SDR της ενεργοποίησης των 200K νευρώνων με $w=2000$ ενεργούς, και ο δενδρίτης που έχει συνάψεις με 30 από τους ενεργούς. Υπό την οπτική του δενδρίτη έχουμε το SDR \hat{A} με μέγεθος 200K και $\hat{w} = 30$, γιατί οι υπόλοιποι ενεργοί νευρώνες δεν τον ερεθίζουν. Αν το όριο ενεργοποίησης του δενδρίτη είναι $\theta=20$, τότε ακόμα και με 30% θόρυβο τυχαίας αναστροφής bit στον πληθυσμό των 200K νευρώνων αναμένουμε ο δενδρίτης να δεχθεί $\hat{w}_n = 70\%\hat{w} = 21 > \theta$ ερεθίσματα και επομένως να ενεργοποιηθεί, παρόλο το θόρυβο.

Θεωρήσαμε ότι το δεύτερο SDR είναι αποτέλεσμα θορύβου στο πρώτο κι επομένως ήταν θεμιτό ο δενδρίτης να τα ταυτίσει. Πόση όμως είναι η πιθανότητα να ενεργοποιηθεί ο ίδιος δενδρίτης από ένα διαφορετικό, τυχαίο SDR με πληθάριθμο \hat{w} που δεν προκύπτει από το αρχικό; Σε αυτήν την περίπτωση η ενεργοποίηση θα ταύτιζε ψευδώς το τυχαίο SDR με το πρώτο. Η πιθανότητα ψευδούς ταύτισης είναι:

$$p_{fp} = \frac{\|overlap_set(\theta)\|}{\text{χωρητικότητα}(N, w)} = \frac{\sum_{b=\theta}^{\hat{w}} \binom{\hat{w}}{b} \binom{N-\hat{w}}{\hat{w}-b}}{\binom{N}{w}} = \frac{8e63}{1e6866} = 6e-6803 \quad (2.1)$$

Καταδεικνύεται έτσι ότι η αναγνώριση SDR είναι μια διαδικασία ανθεκτική στο θόρυβο.

Σύνολα SDR

Διαφάνεια 10.

Επισημαίνεται μονάχα η *σχεδιαστική ελευθερία* που παρέχουν τα SDR. Όσο μεγαλύτερα SDR χρησιμοποιούνται, τόσο αυξάνεται ο χώρος των ενδεχομένων και τόσο πιο απίθανα γίνονται τα λανθασμένα ταιριάσματα.

2.2.3 Μοντέλο δικτύου

Και μόνο ο παραπάνω ανάδρομος ορισμός του νευρώνα παραπέμπει στο ρόλο του ως μονάδα δικτύου.

Η πρώτη οργανωτική δομή νευρώνων, τόσο στο neocortex, όσο και στο HTM, είναι το minicolumn. Συνοπτικά, ο ρόλος των minicolumns είναι να επιτρέπουν διαφορετικές εσωτερικές αναπαραστάσεις των εξωτερικών (feedforward) ερεθισμάτων, ανάλογα με τα συμφραζόμενα. Όλοι οι νευρώνες του minicolumn μοιράζονται *το ίδιο receptive field*, ενώ μεταξύ τους υπάρχει *inhibition*. Έτσι υλοποιείται μια αρχιτεκτονική “winner-takes-all”, όπου, αν κάποιοι νευρώνες είναι ήδη σε predictive κατάσταση όταν έρθει feedforward είσοδος που ερεθίζει το minicolumn, θα ενεργοποιηθούν πρώτοι και θα αποτρέψουν την ενεργοποίηση των υπολοίπων. Η παραπάνω περιγραφή βέβαια προτρέπει λίγο.

Πολλά minicolumns συνθέτουν ένα cortical layer, που αποτελεί τη λειτουργική μονάδα ενός συστήματος HTM. Κάθε minicolumn του cortical layer έχει διαφορετικές feedforward συνδέσεις προς την είσοδο του layer, ενώ διαφορετικά minicolumns μοιράζονται μεταξύ τους συνδέσεις που φέρουν πληροφορίες συμφραζομένων. Σε πολλές περιπτώσεις μπορεί να υπάρχει ολικό inhibition μεταξύ των minicolumns. Αν βέβαια το cortical layer έχει κάποια τοπολογική διάταξη προς την είσοδό του¹, όπως συχνά συμβαίνει στον εγκεφαλικό φλοιό, τότε οι inhibitory συνδέσεις είναι τοπικές και ένα minicolumn δεν ανταγωνίζεται κάποιο άλλο που βρίσκεται μακριά του. Τα λειτουργικά παραδείγματα παρακάτω αποτελούνται από 1 cortical layer. Ένα πιο σύνθετο σύστημα θα μπορούσε να ενώσει εν σειρά πολλά cortical layers.

Οι συνάψεις των HTM νευρώνων είναι επίσης πιστότερες στη βιολογία. Δε χρησιμοποιούν κάποιο βάρος που πολλαπλασιάζει την είσοδο, αλλά είναι δισταθείς, δηλαδή υπάρχουν ή όχι. Ένα μέγεθος όμως που χαρακτηρίζει κάθε σύναψη είναι η *μονιμότητά* της, που διαμορφώνεται δυναμικά από τη μαθησιακή διαδικασία (στην οποία δε θα αναφερθούμε παραπάνω εδώ) με αλγόριθμο τύπου Spike Timing Dependent Plasticity (STDP), που μπορεί ουσιαστικά να εμφανίσει και να εξαφανίσει υποθάλπουσες συνάψεις για να προσαρμόσει το σύστημα στα δεδομένα. Η μονιμότητα μιας σύναψης συγκρίνεται με ένα κατώφλι και, αν είναι πιο πάνω είναι ενεργή, αν είναι πιο κάτω είναι ανενεργή. Μια σύναψη με μονιμότητα λίγο πάνω από το κατώφλι μπορεί δηλαδή να απενεργοποιηθεί εύκολα.

2.2.4 Encoders

Ο εγκέφαλος έχει ως βασική διεπαφή με τον έξω κόσμο τον αμφιβληστροειδή χιτώνα του ματιού και τον κοχλία του αυτιού. Και τα 2 αυτά συστήματα μπορούμε να θεωρήσουμε ότι μετατρέπουν τα φυσικά ερεθίσματα του κόσμου σε μια πρώτη

¹Δες το βίντεο “Topology (Episode 10)” της σειράς “HTM School” [15]

μορφή SDR, για να ερεθίσουν με τη σειρά τους το επόμενο επίπεδο νευρώνων. Το ρόλο αυτό στα συστήματα HTM καλύπτουν οι encoders.

Οι encoders μπορεί να είναι από πολύ απλά έως εξαιρετικά πολύπλοκα συστήματα (βλέπε αμφιβληστροειδή). Στη διαφάνεια 13 δίνεται το παράδειγμα ενός απλού encoder που κωδικοποιεί ακέραιους αριθμούς συγκεκριμένου εύρους. Μπορεί όμως να σχεδιαστεί και encoder που να κωδικοποιεί λέξεις ή και κείμενα, όπως κάνει η εταιρία cortical.io **semantic** στη διαφάνεια 14.

2.2.5 Decoders

2.3 Βασικές λειτουργίες HTM

Παρουσιάζονται ένα σύστημα που κανονικοποιεί τα SDR εισόδου, ο *Spatial Pooler*, και ένα σύστημα που μαθαίνει να αναγνωρίζει ακολουθίες απλών ερεθισμάτων, δηλαδή μια *μνήμη ακολουθιών*. Για το πρώτο προτείνονται τα σχετικά βίντεο από τη σειρά “HTM School” στο youtube, ενώ για το δεύτερο η αναφορά [2]. Λίγα λόγια για τη μνήμη ακολουθιών:

2.3.1 Sequence Memory

Στη *διαφάνεια 17* [10] βλέπουμε ένα cortical layer που αποτελείται από από minicolumns. Στη μέση (C) βλέπουμε πώς αντιδρά το σύστημα καθώς δέχεται για πρώτη φορά τις αλληλουχίες “ABCD” και “XBCY”. Παρατηρούμε ότι κάθε σύμβολο εισόδου ενεργοποιεί συγκεκριμένα minicolumns (αυτά που διεγείρονται περισσότερο), και ότι σε κάθε τέτοιο minicolumn ενεργοποιούνται *όλοι* οι νευρώνες.

Ας επικεντρωθούμε στο τι συμβαίνει όταν λαμβάνεται η είσοδος “A”. Ενεργοποιούνται τα minicolumns που φαίνονται, αλλά κάποιοι από τους ενεργοποιημένους νευρώνες έχουν context συνάψεις (ενεργές ή μη) με κάποιους νευρώνες που ενεργοποιεί η είσοδος “B”. Αυτό φαίνεται στη διαφάνεια 18 A. Αφού το σύστημα δει την αλληλουχία A → B μερικές φορές, συνάψεις αυτού του τύπου ενεργοποιούνται μέσω της διαδικασίας εκμάθησης, ακόμα κι αν στην αρχή ήταν ανενεργές. Έτσι, όταν ξαναέρθει η είσοδος “A”, συγκεκριμένοι νευρώνες των στηλών που αντιστοιχούν στο “B” τίθενται σε predictive state. Όταν λοιπόν όντως έρχεται η είσοδος “B” αμέσως μετά αυτοί οι νευρώνες ενεργοποιούνται πρώτοι και μπλοκάρουν μέσω inhibitory συνδέσεων τους υπόλοιπους του ιδίου minicolumn. Έτσι καταλήγουμε στην ακολουθία της κάτω σειράς (D) της διαφάνειας 17. Τα παραπάνω σημαίνουν ότι το σύστημα *προβλέπει την επόμενη είσοδό του*, που είναι ο βασικός στόχος του προβλήματος εκμάθησης ακολουθιών.

Ο μηχανισμός αυτός οδηγεί σε δύο παρατηρήσεις: η εσωτερική αναπαράσταση κάθε συμβόλου είναι *διαφορετική ανάλογα με τα συμφραζόμενα* και το σύστημα μπορεί να πραγματοποιεί *πολλαπλές προβλέψεις ταυτόχρονα*.

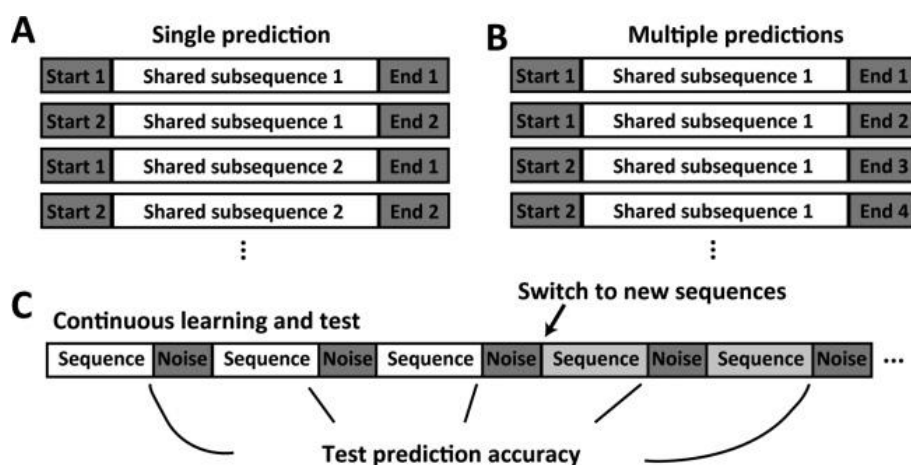
Το πρώτο φαίνεται στην κάτω σειρά (D) από τους διαφορετικούς νευρώνες που ενεργοποιούνται για να συμβολίσουν την ίδια είσοδο “B”, ανάλογα αν το προηγούμενο σύμβολο ήταν “A” ή “X”, οπότε προκύπτει η εσωτερική αναπαράσταση B’ ή B” αντίστοιχα. Το δεύτερο σημείο φαίνεται στη διαφάνεια 18 B, όπου δίνεται η είσοδος “B”, χωρίς να έχει προηγηθεί “A” ή “X”. Το σύστημα τότε προβλέπει την ένωση των 2 καταστάσεων C’ και C”, οι οποίες ενεργοποιούνται παράλληλα όταν όντως έρχεται η είσοδος C. Εκείνη τη στιγμή λοιπόν το σύστημα σωστά προβλέπει ότι η συνέχεια της ακολουθίας θα είναι ή “D” ή “Y”.

2.4 Αποτελέσματα - Προσομοιώσεις

Το HTM δοκιμάστηκε τόσο σε τεχνητά, όσο και σε πραγματικά δεδομένα για να αξιολογηθεί η δυνατότητα του να παρέχει ακριβείς προβλέψεις. Παρακάτω θα αναλυθούν δύο παραδείγματα και θα δοθούν τα συγκριτικά διαγράμματα των επιδόσεων με άλλα νευρωνικά δίκτυα.

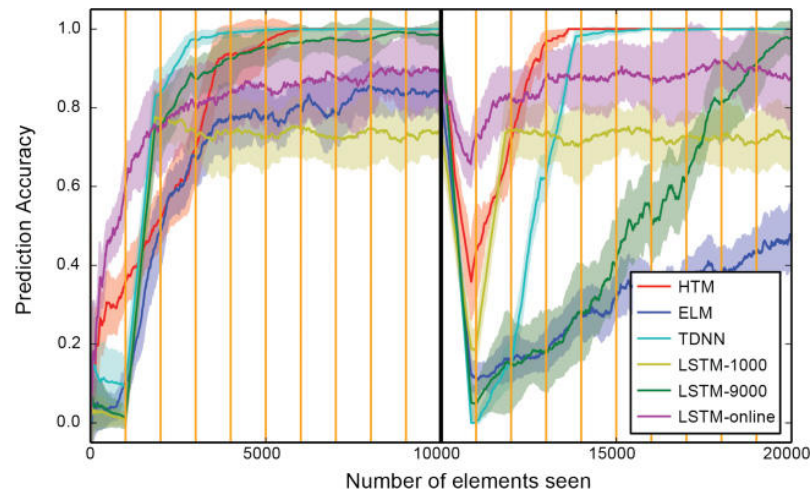
Πρόβλεψη συμβόλου ακολουθίας

Οι αλγόριθμοι του Hierarchical Temporal Memory (HTM) δοκιμάστηκαν σε διαφορετικά προβλήματα [9], [10], ώστε να αξιολογηθεί η επίδοση του συστήματος σε σχέση με άλλες υλοποιήσεις. Στο πρώτο παράδειγμα έχουν δημιουργηθεί δύο ακολουθίες από σύμβολα, σύμφωνα με την εικόνα (2.2). Στόχος είναι η πρόβλεψη του τελευταίου συμβόλου.



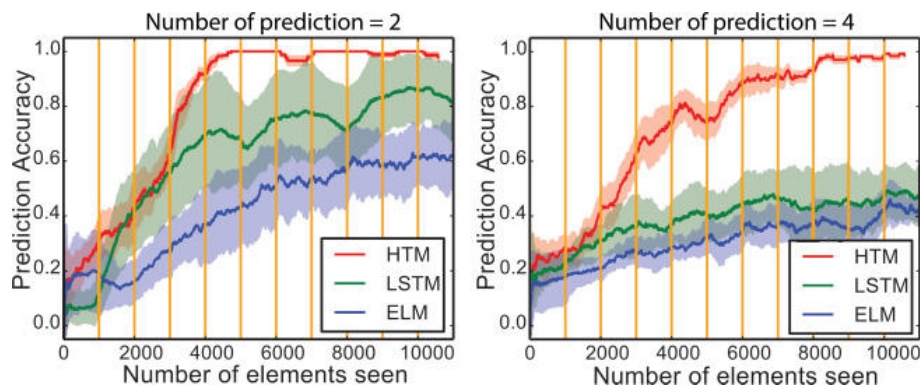
Σχήμα 2.2: Ακολουθίες συμβόλων για την προσομοίωση

Οι ακολουθίες φτιάχτηκαν έτσι ώστε να χρειάζεται ένα βάθος μνήμης τουλάχιστον 2 παρελθοντικών συμβόλων για την επιτυχή πρόβλεψη. Το πρώτο set ακολουθιών έχει δημιουργηθεί για να ελεγχθεί η δυνατότητα απλής πρόβλεψης, καθώς κάθε ακολουθία προσδιορίζεται πλήρως από τα πεδία “Start” και “Shared Subsequence”. Αντιθέτως, στο δεύτερο set δύο ακολουθίες μπορεί να έχουν ίδια τα παραπάνω πεδία. Σε αυτή την περίπτωση το νευρωνικό δίκτυο θα πρέπει να επιτύχει την ορθή πρόβλεψη όλων των πιθανών καταλήξεων, δηλαδή την ταυτόχρονη πολλαπλή πρόβλεψη. Τέλος, ανάμεσα στις ακολουθίες προστέθηκε ένα σύμβολο θορύβου, ώστε να εξεταστεί η ανθεκτικότητα του συστήματος σε αυτόν.



Σχήμα 2.3: Απόδοση διαφορετικών υλοποιήσεων για απλή πρόβλεψη

Στο σχήμα (2.3) δίνεται η συγκριτική απόδοση των διαφορετικών υλοποιήσεων νευρωνικών δικτύων για το πρώτο set ακολουθιών. Τα δίκτυα που υλοποιούνται είναι το HTM, το ELM, το TDNN και το LSTM. Ο προσδιορισμός στο LSTM αφορά το πλήθος των δειγμάτων που χρησιμοποιούνται για retraining σε κάθε κατακόρυφη κίτρινη γραμμή του διαγράμματος. Στο στοιχείο 10000 που σημειώνεται με κάθετη κατακόρυφη γραμμή, αλλάζουν αμοιβαία τα δύο τελευταία σύμβολα κάθε ακολουθίας και μελετάται η προσαρμογή του δικτύου στο νέο input space. Βλέπουμε ότι το HTM πετυχαίνει αρχικά σε έναν ικανοποιητικό αριθμό δειγμάτων ακρίβεια 100%. Το πιο σημαντικό είναι, βέβαια, ότι μέσω online training αναπροσαρμόζεται ταχύτερα από κάθε άλλη υλοποίηση στο νέο χώρο ακολουθιών εισόδου.

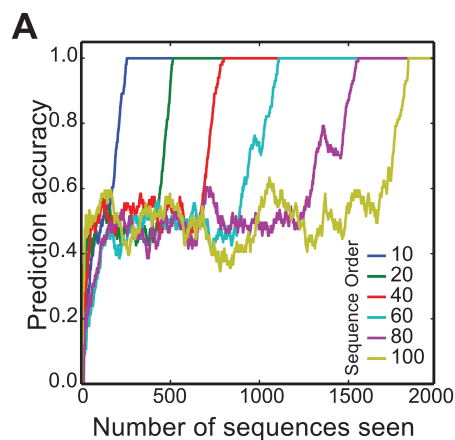


Σχήμα 2.4: Απόδοση δικτύων για πολλαπλές προβλέψεις

Οι πολλαπλές προβλέψεις εξετάζονται από το δεύτερο set και τα συγκριτικά αποτελέσματα για διπλή και τετραπλή πρόβλεψη δίνεται στο σχήμα (2.4). Όπως φαίνεται, το HTM είναι το μοναδικό σύστημα που πετυχαίνει μετά από κάποιο διάστημα ακρίβεια σχεδόν 100%. Αυτό απορρέει από την αναπαράσταση μέσω SDR και τη δυνατότητα του Temporal Pooler για πολλαπλές προβλέψεις.

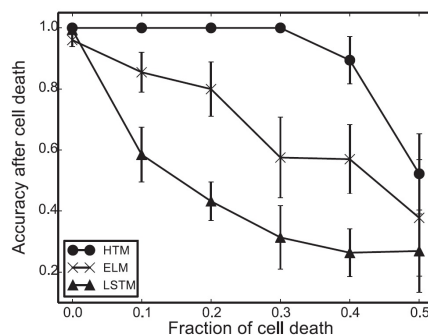
Στη συνέχεια, μελετήθηκε η απόδοση του HTM για ακολουθίες διαφορετικού μήκους. Ένα από τα σημαντικότερα στοιχεία του Temporal Pooler είναι η δυνατότητα αναπροσαρμογής online του «βάθους» μνήμης που χρησιμοποιείται για την πρόβλεψη. Έτσι, όπως παρατηρούμε και στο διάγραμμα (2.5), πετυχαίνει πρόβλεψη ακόμα και τάξης 100, που σημαίνει ότι απαιτείται μνήμη 99 συμβόλων για να προβλεφθεί επιτυχώς το τελευταίο. Το πλήθος των ακολουθιών που απαιτείται για να εκπαιδευτεί

το σύστημα και να πετύχει ακρίβεια κοντά στο 100% αυξάνει γραμμικά με την τάξη της ακολουθίας.



Σχήμα 2.5: Απόδοση HTM συναρτήσει της τάξης της ακολουθίας

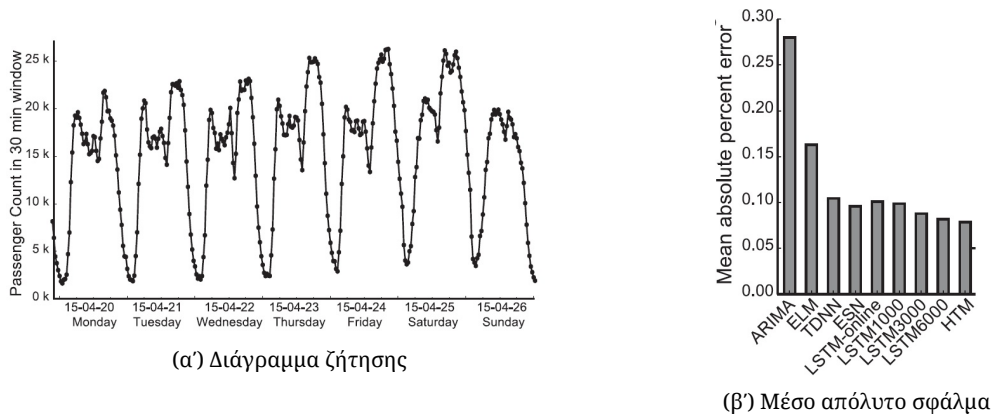
Τέλος, τα διαφορετικά συστήματα εξετάστηκαν και ως προς την ανθεκτικότητα σε καταστροφή του δικτύου μέσω νεκρών νευρώνων. Το HTM έμεινε ανεπηρέαστο ακόμα για 30% νεκρούς νευρώνες, συνθήκες κατά τις οποίες το ELM και το LSTM είχαν υποβαθμιστεί αρκετά.



Σχήμα 2.6: Ανθεκτικότητα υλοποιήσεων σε νεκρούς νευρώνες

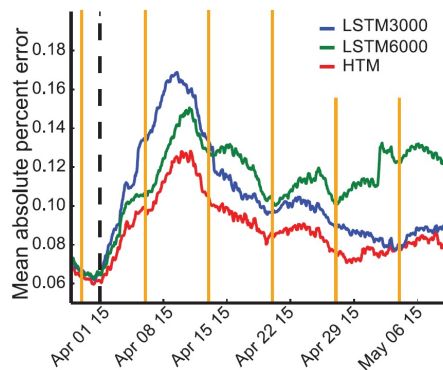
Πρόβλεψη της ζήτησης taxi στη Νέα Υόρκη

Το Hierarchical Temporal Memory δοκιμάστηκε και σε πραγματικά streaming δεδομένα. Πιο συγκεκριμένα, ζητήθηκε η πρόβλεψη της ζήτησης στα taxi της Νέας Υόρκης 2.5 ώρες πριν την πραγματική μέτρηση. Η ζήτηση ποσοτικοποιήθηκε ως η συνολική εξυπηρέτηση πελατών σε ένα χρονικό παράθυρο 30 λεπτών. Εξετάστηκαν διάφοροι τύποι νευρωνικών δικτύων, καθώς και το στατιστικό μοντέλο ARIMA.



Σχήμα 2.7: Πρόβλεψη ζήτησης των taxi

Στη συνέχεια, τα δεδομένα εισόδου τροποποιήθηκαν τεχνητά ανεβάζοντας ή μειώνοντας τη ζήτηση κατά 20% σε διαφορετικές χρονικές περιόδους εντός κάθε μέρας. Στόχος ήταν να μελετηθεί η προσαρμογή του δικτύου στις νέες (τεχνητές) συνήθειες των επιβατών taxi της Νέας Υόρκης. Η αλλαγή πραγματοποιήθηκε την 1η Απριλίου και εντός 2 βδομάδων το σύστημα HTM αναπροσάρμοσε το μοντέλο του, ώστε να πετυχαίνει εκ νέου σωστές προβλέψεις. Σε αντίθεση, το LSTM είτε απέτυχε να επανέρθει, είτε χρειάστηκε πολύ περισσότερο διάστημα για retraining των παραμέτρων του.



Σχήμα 2.8: Προσαρμογή δικτύων σε μεταβολή της ζήτησης

Βιβλιογραφία

- [1] M. Mohammadi, A. Al-Fuqaha, S. Sorour και M. Guizani, «Deep Learning for IoT Big Data and Streaming Analytics: A Survey,» *IEEE Communications Surveys Tutorials*, τόμ. 20, αρθμ. 4, σσ. 2923–2960, Fourthquarter 2018, ISSN: 1553-877X. doi: 10.1109/COMST.2018.2844341.
- [2] P. Kanerva, «Associative Neural Memories,» στο, M. H. Hassoun, επιμελητής, New York, NY, USA: Oxford University Press, Inc., 1993, κεφ. Sparse Distributed Memory and Related Models, σσ. 50–76, ISBN: 0-19-507682-6.
- [3] V. B. Mountcastle, «The columnar organization of the neocortex,» *Brain: A Journal of Neurology*, τόμ. 120 (Pt 4), σσ. 701–722, Απρ. 1997, ISSN: 0006-8950. doi: 10.1093/brain/120.4.701. pmid: 9153131.
- [4] J. C. Horton και D. L. Adams, «The Cortical Column: A Structure without a Function,» *Philosophical Transactions of the Royal Society B: Biological Sciences*, τόμ. 360, αρθμ. 1456, σσ. 837–862, 29 Απρ. 2005, ISSN: 0962-8436. doi: 10.1098/rstb.2005.1623. pmid: 15937015.
- [5] T. Freund και S. Kali, «Interneurons,» *Scholarpedia*, τόμ. 3, αρθμ. 9, σ. 4720, 1 Σεπτ. 2008, ISSN: 1941-6016. doi: 10.4249/scholarpedia.4720.
- [6] P. Wang, «What Do You Mean by “AI”?,» **presented at** Proceedings of the 2008 Conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference, IOS Press, 20 Ιούν. 2008, σσ. 362–373, ISBN: 978-1-58603-833-5. διεύθυν.: <http://dl.acm.org/citation.cfm?id=1566174>. 1566207 (επίσκεψη 28/05/2019).
- [7] J. Defelipe, H. Markram και K. S. Rockland, «The Neocortical Column,» *Frontiers in Neuroanatomy*, τόμ. 6, 2012, ISSN: 1662-5129. doi: 10.3389/fnana.2012.00022.
- [8] Y. Bengio, D.-H. Lee, J. Bornschein, T. Mesnard και Z. Lin, «Towards Biologically Plausible Deep Learning,» 13 Φεβ. 2015. arXiv: 1502.04156 [cs]. διεύθυν.: <http://arxiv.org/abs/1502.04156> (επίσκεψη 28/05/2019).
- [9] A. Lavin και S. Ahmad, «Evaluating Real-Time Anomaly Detection Algorithms – The Numenta Anomaly Benchmark,» στο *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, Δεκ. 2015, σσ. 38–44. doi: 10.1109/ICMLA.2015.141.
- [10] Y. Cui, S. Ahmad και J. Hawkins, «Continuous online sequence learning with an unsupervised neural network model,» *Neural Computation*, 2016.
- [11] P. Haueis, «The Life of the Cortical Column: Opening the Domain of Functional Architecture of the Cortex (1955–1981),» *History and Philosophy of the Life Sciences*, τόμ. 38, αρθμ. 3, 2016, ISSN: 0391-9714. doi: 10.1007/s40656-016-0103-4. pmid: 27325058.

- [12] J. Hawkins και S. Ahmad, «Why Neurons Have Thousands of Synapses, a Theory of Sequence Memory in Neocortex,» *Frontiers in Neural Circuits*, τόμ. 10, 2016, ISSN: 1662-5110. DOI: 10.3389/fncir.2016.00023.
- [13] —, «Why Neurons Have Thousands of Synapses, a Theory of Sequence Memory in Neocortex,» *Frontiers in Neural Circuits*, τόμ. 10, σ. 23, 2016, ISSN: 1662-5110. DOI: 10.3389/fncir.2016.00023. Διεύθυν.: <http://journal.frontiersin.org/article/10.3389/fncir.2016.00023>.
- [14] Y. LeCun, «Deep learning,» CERN talk, 2016, Διεύθυν.: <https://indico.cern.ch/event/510372/attachments/1245509/1840815/lecun-20160324-cern.pdf>.
- [15] M. Taylor. (2016). HTM School. Video lectures on YouTube, Διεύθυν.: <https://www.youtube.com/playlist?list=PL3yXMgtrZmDqhsFQzwUC9V8MeeV0Q7eZ9>.
- [16] C. M. Vineyard και S. J. Verzi, «Overcoming the Static Learning Bottleneck - the need for adaptive neural learning,» στο *2016 IEEE International Conference on Rebooting Computing (ICRC)*, Οκτ. 2016, σσ. 1–3. DOI: 10.1109/ICRC.2016.7738692.
- [17] S. Sabour, N. Frosst και G. E. Hinton, «Dynamic Routing Between Capsules,» στο *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan και R. Garnett, επιμελητές, Curran Associates, Inc., 2017, σσ. 3856–3866. Διεύθυν.: <http://papers.nips.cc/paper/6975-dynamic-routing-between-capsules.pdf> (επίσκεψη 30/05/2019).
- [18] V. Losing, B. Hammer και H. Wersing, «Incremental On-Line Learning: A Review and Comparison of State of the Art Algorithms,» *Neurocomputing*, τόμ. 275, σσ. 1261–1274, 31 Ιαν. 2018, ISSN: 0925-2312. DOI: 10.1016/j.neucom.2017.06.084.
- [19] O. Nachum, S. (Gu, H. Lee και S. Levine, «Data-Efficient Hierarchical Reinforcement Learning,» στο *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi και R. Garnett, επιμελητές, Curran Associates, Inc., 2018, σσ. 3303–3313. Διεύθυν.: <http://papers.nips.cc/paper/7591-data-efficient-hierarchical-reinforcement-learning.pdf> (επίσκεψη 30/05/2019).
- [20] S. Sabour, N. Frosst και G. Hinton, «Matrix Capsules with EM Routing,» **presented at** 6th International Conference on Learning Representations, ICLR, 2018.
- [21] P. Xiong, Y. Zhu, Z. Sun, Z. Cao, M. Wang, Y. Zheng, J. Hou, T. Huang και Z. Que, «Application of Transfer Learning in Continuous Time Series for Anomaly Detection in Commercial Aircraft Flight Data,» στο *2018 IEEE International Conference on Smart Cloud (SmartCloud)*, Σεπτ. 2018, σσ. 13–18. DOI: 10.1109/SmartCloud.2018.00011.
- [22] O. Vinyals, I. Babuschkin, J. Chung, M. Mathieu και D. Silver. (2019). AlphaStar: Mastering the Real-Time Strategy Game StarCraft II, Διεύθυν.: <https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/>.
- [23] Fabuio. (2017-07-06). Pyramidal neuron, Διεύθυν.: https://upload.wikimedia.org/wikipedia/commons/c/c1/Piramidal_cell.svg (επίσκεψη 31/05/2019).