



**ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ  
«ΠΡΟΗΓΜΕΝΑ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ»**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**Μηχανές Μάθησης Υψηλής Απόδοσης για Προβλήματα  
Κατηγοριοποίησης - Εφαρμογή στην Ταξινόμηση Κτιρίων σε  
Ενεργειακές Κλάσεις**

**Άννα Ν. Κόκκινου**

**Επιβλέποντες: Παναγιώτης Σταματόπουλος, Επίκουρος Καθηγητής**

**ΑΘΗΝΑ**

**ΜΑΡΤΙΟΣ 2015**

## **ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

Οι μηχανές μάθησης υψηλής απόδοσης σε προβλήματα κατηγοριοποίησης

**Άννα Ν. Κόκκινου**

**A.M.: 1253**

**ΕΠΙΒΛΕΠΟΝΤΕΣ:** Παναγιώτης Σταματόπουλος, Επίκουρος Καθηγητής

**ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:** Βασίλειος Ζησιμόπουλος, Καθηγητής

Μάρτιος 2015

## ΠΕΡΙΛΗΨΗ

Στην εργασία αυτή μελετάται και αναλύεται ο αλγόριθμος Extreme Learning Machine, καθώς και η εφαρμογή του σε προβλήματα κατηγοριοποίησης και ταξινόμησης.

Τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης, χρησιμοποιούνται κατά κόρον σε προβλήματα ταξινόμησης, όμως η διαδικασία εκπαίδευσης τους με τους παραδοσιακούς αλγόριθμους είναι εξαιρετικά αργή, πράγμα που οφείλεται στην εκτεταμένη χρήση των αλγορίθμων που βασίζονται στην κλίση και στο ότι οι παράμετροι των δικτύων υπολογίζονται σε κάθε επανάληψη.

Ο αλγόριθμος ELM διαφοροποιείται συγκριτικά με τους παραδοσιακούς αλγόριθμους μάθησης στο ότι είναι εξαιρετικά γρήγορος, επιτυγχάνοντας ταυτόχρονα μεγάλη ακρίβεια στην ταξινόμηση. Επιπλέον, οι παράμετροι του δικτύου επιλέγονται τυχαία και δεν επαναυπολογίζονται ξανά ως το πέρας της διαδικασίας εκπαίδευσης.

Παρουσιάζονται αρκετές παραλλαγές του αλγορίθμου ELM, με τις οποίες ξεπερνιούνται κάποιες δυσκολίες και μειονεκτήματα του παραδοσιακού αλγορίθμου και επιτυγχάνονται ακόμα καλύτεροι χρόνοι και ταξινόμηση, καθώς και κάποια υβριδικά μοντέλα συνδυασμού του κλασικού αλγορίθμου ELM με κάποιον παραδοσιακό αλγόριθμο μάθησης.

Επιπρόσθετα, αναφέρονται ενδεικτικά κάποιες εφαρμογές του αλγορίθμου ELM σε προβλήματα ταξινόμησης που βρίσκουν εφαρμογή σε πολλαπλά επιστημονικά πεδία, όπως στη βιοιατρική, στη μηχανική, στη βιολογία, καθώς και στη μουσική.

Στη συνέχεια, αναλύεται ένα πρόβλημα ταξινόμησης, που αφορά στην ταξινόμηση κτιρίων στην ενεργειακή τους κλάση ανάλογα με το φορτίο θέρμανσης και ψύξης που χρειάζονται. Περιγράφεται αναλυτικά πώς προκύπτει η συνολική πρωτογενής καταναλισκόμενη ενέργεια από τα φορτία θέρμανσης και ψύξης και στη συνέχεια με ποια μεθοδολογία προκύπτει η ενεργειακή κλάση για το κάθε κτίριο.

Τέλος, γίνεται πειραματική αξιολόγηση του αλγορίθμου για το συγκεκριμένο πρόβλημα ταξινόμησης, συγκρίνοντας τον με άλλους αλγορίθμους κατηγοριοποίησης δεδομένων.

**ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ:** Κατηγοριοποίηση Δεδομένων

**ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ:** αλγόριθμοι εκπαίδευσης, ταξινόμηση, ενεργειακή κλάση, συνολική καταναλισκόμενη πρωτογενής ενέργεια, φορτίο θέρμανσης, φορτίο ψύξης

## **ABSTRACT**

The topic of this thesis is the study and analysis of the extreme Learning Machine Algorithm and its application on classification problems.

Feed forward neural networks have been extensively used on classification problems, but their training time with traditional learning algorithms is extremely slow. Two key reasons for that may be: the slow gradient-based learning algorithms are extensively used to train neural networks, and all the parameters of the networks are tuned iteratively by using such learning algorithms.

The ELM algorithm differs compared to traditional learning algorithms because it is extremely fast, while achieving high accuracy in classification. In addition, the parameters of the network are selected randomly and not recalculated until the end of the training process.

Several ELM variants are presented, which manage to overcome the difficulties and disadvantages of the classical algorithm achieving better training times and training accuracy. Some hybrid models are presented as well combining the classic algorithm ELM with a traditional learning algorithm.

Additionally, some ELM algorithm's applications are mentioned in classification problems in multiple scientific fields, such as biomedical, engineering, biology and music.

Then, a new classification problem is analyzed, concerning the classification of buildings in an energy class depending on required heating load and cooling load. It is elaborated how the total primary energy of every building comes out from heating load and cooling load and there is a thorough description of the methodology used for buildings classification regarding this total energy amount.

Finally experimental results are presented and Extreme Learning Machine Algorithm is compared with other learning algorithms.

**SUBJECT AREA:** Classification Problems

**KEYWORDS:** learning algorithm, classification, energy class, total primary energy, heating load, cooling load

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Κατά την εκπόνηση της παρούσας διπλωματικής μελέτης, αρωγός μου υπήρξε η ενεργός παρουσία του επιβλέποντος καθηγητή μου κ. Παναγιώτη Σταματόπουλου, καθώς με την άμεση ανταπόκρισή του, την υπομονή και την κατανόησή του με βοήθησε να την ολοκληρώσω με επιτυχία.

Επίσης είμαι ευγνώμων στον αξιολογητή της διπλωματικής μου εργασίας κ. Ζησιμόπουλο για την προσεκτική ανάγνωση της εργασίας μου και τις χρήσιμες υποδείξεις του.

Τέλος η πολύτιμη συμπαράσταση του Κωνσταντίνου Γραμματικού, ο οποίος με χαρά συνέβαλε στην επιμέλεια και τον έλεγχο της διπλωματικής μου εργασίας.

Θερμές ευχαριστίες.

# ΠΕΡΙΕΧΟΜΕΝΑ

<b>1. Εισαγωγή.....</b>	<b>11</b>
<b>2. Τεχνητοί Νευρώνες .....</b>	<b>12</b>
<b>3. Extreme Learning Machine .....</b>	<b>14</b>
<b>4. Παραλλαγές του αλγορίθμου Extreme Learning Machine .....</b>	<b>19</b>
4.1 E-ELM (Evolutionary ELM) .....	19
4.2 C-ELM (Complex ELM) .....	20
4.3 OS - ELM / EOS-ELM / OS - Fuzzy - ELM (Online Sequential / Enhanced Online Sequential / On Line Sequential Fuzzy ELM) .....	21
4.4 SCELM (Sine / Cosine ELM) .....	23
4.5 I-ELM/CI-ELM / EI-ELM (Incremental / Convex Incremental /Enhanced Incremental ELM) .....	24
4.6 OP-ELM / P-ELM (Optimally Pruned / Pruned ELM) .....	25
4.7 R-ELM (Regularized ELM) .....	26
4.8 SB-ELM (Sparse-Bayesian ELM).....	26
4.9 S-ELM (Sparse ELM) .....	27
<b>5. Υβριδικά μοντέλα .....</b>	<b>29</b>
5.1 Υβριδικό ELM-SVM.....	29
5.2 Υβριδικό ELM- GMC .....	30
<b>6. Εφαρμογές του αλγόριθμου ELM σε προβλήματα ταξινόμησης.....</b>	<b>33</b>
6.1 Μόρια συμπλέγματος ιστοσυμβατότητας.....	33
6.2 Ταξινόμηση ψυχικών καταστάσεων από ηλεκτροσήματα του εγκεφάλου .....	33
6.3 Ταξινόμηση με ηλεκτρονική μύτη .....	35
6.4 Ταξινόμηση πρωτεϊνών .....	35
6.5 Ταξινόμηση ειδών μουσικής .....	36
6.6 Αναγνώριση προσώπων - Ταξινόμηση αντικειμένων -Αναγνώριση δραστηριοτήτων .....	38
6.6.1 Πρόσωπα.....	38
6.6.2 Αντικείμενα .....	40
6.6.3 Δραστηριότητες .....	40
6.7 Ταξινόμηση μικροσυστοιχιών γονιδίων για τη διάγνωση καρκίνου .....	41
<b>7. Παρουσίαση - ανάλυση προβλήματος ταξινόμησης κτιρίων σε ενεργειακές κλάσεις.....</b>	<b>42</b>
7.1 Εισαγωγή.....	42

7.1.1 Ορισμοί.....	43
7.1.2 Συνοπτική παρουσίαση του προβλήματος.....	43
<b>7.2 Dataset (από UCI) .....</b>	<b>44</b>
<b>7.3 Πρόβλημα ταξινόμησης ενεργειακών κλάσεων .....</b>	<b>46</b>
7.3.1 Κτίριο Αναφοράς .....	47
7.3.2 Βαθμομέρες και βαθμώρες .....	47
7.3.2.1 Βαθμομέρες θέρμανσης.....	47
7.3.2.2 Βαθμώρες ψύξης.....	48
7.3.3 Κλιματικές ζώνες .....	50
7.3.4 Ενεργειακές κλάσεις.....	51
7.3.5 Παραδοχές.....	52
<b>8. Πειράματα.....</b>	<b>53</b>
8.1 Πειράματα με τον αλγόριθμο ELM .....	57
8.2 Πειράματα με άλλους αλγορίθμους .....	57
<b>9. Αποτελέσματα – Συμπεράσματα .....</b>	<b>60</b>
<b>ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ .....</b>	<b>62</b>
<b>ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ .....</b>	<b>63</b>
<b>ΑΝΑΦΟΡΕΣ.....</b>	<b>65</b>

## ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1: Δείγμα του dataset από το UCI .....	45
Σχήμα 2: Δείγμα του αρχείου excel με τους αναλυτικούς υπολογισμούς για τις ενεργειακές κλάσεις (μέρος 1) .....	53
Σχήμα 3: Δείγμα του αρχείου excel με τους αναλυτικούς υπολογισμούς για τις ενεργειακές κλάσεις (μέρος 2) .....	54
Σχήμα 4: Βαθμοημέρες θέρμανσης και ψύξης για τις τρεις πόλεις των πειραμάτων .....	54
Σχήμα 5: Σύνοψη αποτελεσμάτων από την εκτέλεση των πειραμάτων .....	60



## ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Γραμμικό μοντέλο τεχνητού νευρώνα .....	12
Εικόνα 2: Νευρωνικό δίκτυο πρόσθιας τροφοδότησης με ένα κρυφό επίπεδο .....	14
Εικόνα 3: Σχηματική απεικόνιση κλιματικών ζωνών ελληνικής επικράτειας.....	51
Εικόνα 4: Εισαγωγή του dataset στο weka.....	58

## ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Βαθμομέρες θέρμανσης DD με θερμοκρασία αναφοράς 18 °C .....	49
Πίνακας 2: Βαθμώρες ψύξης CDH με θερμοκρασία αναφοράς 26 °C.....	50
Πίνακας 3: Κατηγορίες ενεργειακής απόδοσης κτιρίων .....	52

## 1. Εισαγωγή

Μέχρι σήμερα, τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης (*feed forward neural networks*) βρίσκουν εφαρμογή σε πολλά πεδία λόγω της ικανότητας τους να μπορούν να προσεγγίσουν μη γραμμικές αντιστοιχίσεις απευθείας από τα δεδομένα εισόδου και λόγω του ότι μπορούν να παρέχουν αποδοτικά μοντέλα για μία μεγάλη κατηγορία φυσικών και τεχνητών φαινομένων, που είναι δύσκολο να διαχειριστούν από άλλες τεχνικές κατηγοριοποίησης.

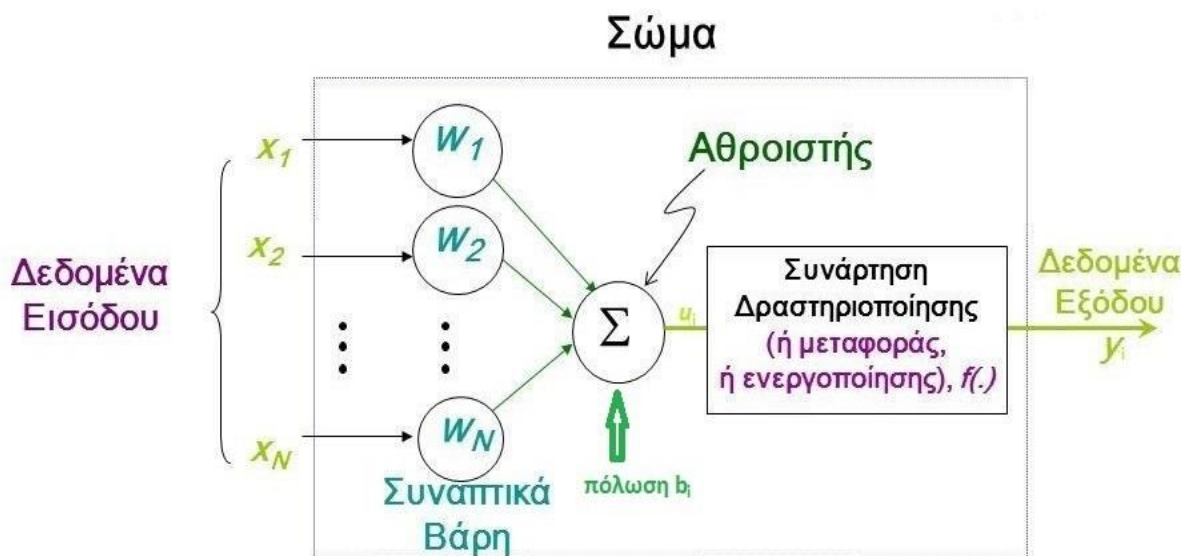
Από την άλλη πλευρά η διαδικασία μάθησης των νευρωνικών δικτύων με βάση τους παραδοσιακούς αλγόριθμους μάθησης είναι εξαιρετικά αργή, πράγμα που οφείλεται στην εκτεταμένη χρήση των αλγορίθμων που βασίζονται στην κλίση (*gradient-based algorithms*) για την εκπαίδευση και στο ότι οι παράμετροι των δικτύων υπολογίζονται σε κάθε επανάληψη.

Πιο συγκεκριμένα, υπάρχει ένα άνω όριο ταχύτητας που οι περισσότεροι (αν όχι όλοι) παραπάνω αλγόριθμοι δεν μπορούν να ξεπεράσουν και συνεπώς ο χρόνος που απαιτείται για την εκπαίδευση ενός νευρωνικού δικτύου είναι πάρα πολύ μεγάλος ακόμα και στην περίπτωση απλών εφαρμογών. Επιπρόσθετα, οι αλγόριθμοι αυτοί λειτουργούν μόνο για διαφορίσιμες συναρτήσεις ενεργοποίησης και έχουν σαν μοναδικό στόχο τους να φτάσουν το σφάλμα της εκπαίδευσης στο ελάχιστο. Αντιμετωπίζουν αρκετές δυσκολίες που σχετίζονται με το τοπικό ελάχιστο (*local minima*), τον ακατάλληλο ρυθμό μάθησης (*learning rate*) την υπερβολική προσαρμογή/ υπερεκπαίδευση (*overfitting*). Για αυτό το λόγο απαιτούνται μέθοδοι εγκυρότητας και οριστικής παύσης των αλγορίθμων αυτών ώστε να μειωθεί η συνάρτηση κόστους. Τέλος αξίζει να σημειωθεί πως οι αλγόριθμοι αυτοί εκπαιδεύουν αποδοτικά νευρωνικά δίκτυα πρόσθιας τροφοδότησης με παραπάνω από ένα κρυφά επίπεδα.

Σε αντίθεση με όλες τις παραπάνω μεθόδους μάθησης το 2004 προτάθηκε ο νέος αλγόριθμος μηχανής μάθησης υψηλής απόδοσης (*Extreme Learning Machine (ELM)*) που εφαρμόζεται κυρίως σε νευρωνικά δίκτυα πρόσθιας τροφοδότησης με ένα κρυφό επίπεδο (*Single Layer Feed forward Networks SLFN*). Ο απώτερος στόχος της συγκεκριμένης πρότασης είναι η καλύτερη δυνατή απόδοση γενίκευσης και η εκπληκτικά γρήγορη ταχύτητα μάθησης.

## 2. Τεχνητοί Νευρώνες

Τα νευρωνικά δίκτυα αποτελούνται από τεχνητούς νευρώνες, οι οποίοι, σε αντιστοιχία με τους βιολογικούς, μοντελοποιούν τον τρόπο με τον οποίο λειτουργεί το νευρικό σύστημα. Ο τεχνητός νευρώνας είναι μονάδα επεξεργασίας πληροφορίας, η οποία είναι στοιχειώδης για τη λειτουργία ενός νευρωνικού δικτύου. Το πιο συνηθισμένο μοντέλο νευρώνα απεικονίζεται στο παρακάτω σχήμα :



Εικόνα 1: Γραμμικό μοντέλο τεχνητού νευρώνα

Το μοντέλο αυτό αποτελείται από τρία βασικά στοιχεία:

- Ένα σύνολο συνάψεων ή διασυνδέσεων, μέσω των οποίων τα δεδομένα εισόδου ( $x_1, x_2, \dots, x_N$ ), πολλαπλασιάζονται με ένα συναπτικό βάρος ( $w_1, w_2, \dots, w_N$ ) αντίστοιχα, πριν καταλήξουν στον κόμβο άθροισης.
- Έναν αθροιστή, ο οποίος αθροίζει τα σταθμισμένα σήματα εισόδου (δηλαδή τα παραπάνω γινόμενα), καθώς και την πόλωση του νευρώνα. Παράγεται έτσι ένα τοπικό πεδίο  $u_k$ , το οποίο δίδεται από την παρακάτω εξίσωση

$$u_i = \sum_{j=1..N} (w_j x_j + b_j) \quad (1)$$

- Μία συνάρτηση ενεργοποίησης (*activation function*), που συμβολίζεται ως  $f(\cdot)$  και είναι ένα είδος φίλτρου που διαμορφώνει την τελική τιμή του σήματος εξόδου  $y_i$ .

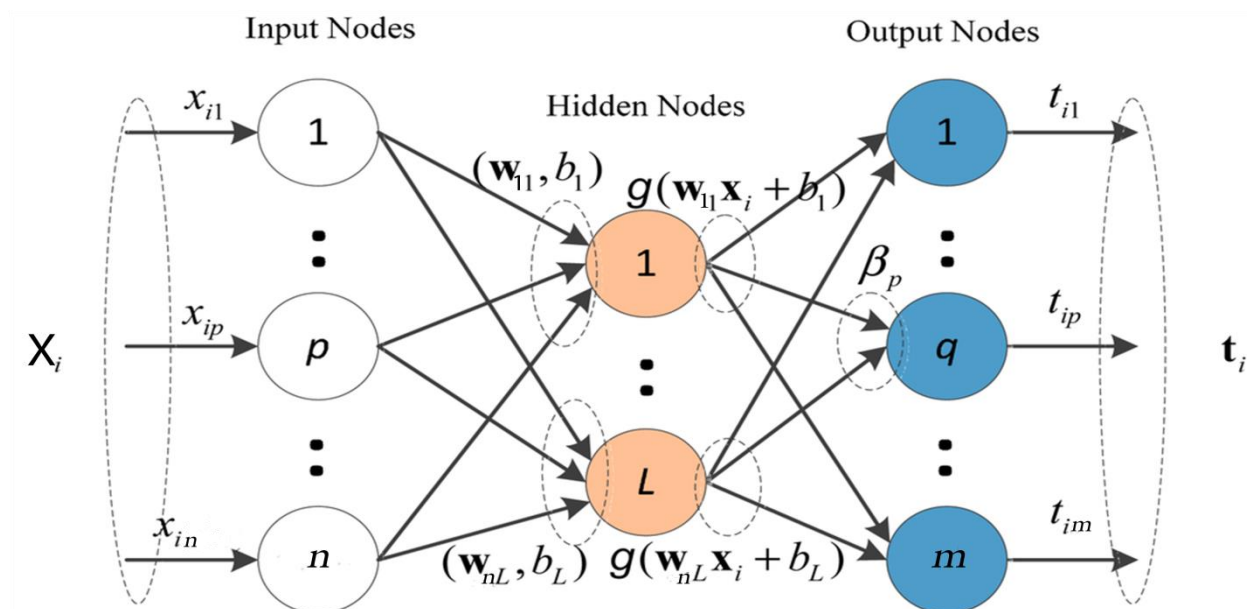
Το πιο απλό νευρωνικό δίκτυο αποτελείται από έναν νευρώνα, συνεπώς η είσοδος και η έξοδος του δικτύου ταυτίζονται με την είσοδο και έξοδο του νευρώνα. Στην περίπτωση που έχουμε περισσότερες από μία εισόδους ή εξόδους το νευρωνικό δίκτυο περιλαμβάνει περισσότερους νευρώνες.

Πολλοί νευρώνες, σαν αυτόν που περιγράφηκε παραπάνω συνδέονται μεταξύ τους βάση συγκεκριμένων αρχιτεκτονικών για τη δημιουργία ενός τεχνητού νευρωνικού δικτύου. Οι αρχιτεκτονικές μπορούν να κατηγοριοποιηθούν σε τρεις κατηγορίες, βάση του αριθμού των κρυφών επιπέδων τους και την παρουσία ή μη ανάδρασης (δίκτυα πρόσθιας τροφοδότησης ενός επιπέδου, πολυεπίπεδα δίκτυα πρόσθιας τροφοδότησης, αναδρομικά δίκτυα).

Λόγω του ότι αντικείμενο της παρούσας αποτελεί ο αλγόριθμος ELM, ο οποίος εφαρμόζεται κατά κόρον σε δίκτυα πρόσθιας τροφοδότησης με ένα κρυφό επίπεδο, παρακάτω αναλύεται μόνο η συγκεκριμένη κατηγορία νευρωνικών δικτύων.

### 3. Extreme Learning Machine

Στην παρακάτω εικόνα φαίνεται ένα δίκτυο πρόσθιας τροφοδότησης με ένα κρυφό επίπεδο:



Εικόνα 2: Νευρωνικό δίκτυο πρόσθιας τροφοδότησης με ένα κρυφό επίπεδο

Τα  $x_{i1}, x_{i2}, \dots, x_{in}$  αναπαριστούν τα σήματα εισόδου (στους κόμβους εισόδου, ο αριθμός των οποίων είναι  $n$ , δεν εκτελείται κάποιος υπολογισμός), τα  $t_{i1}, t_{i2}, \dots, t_{im}$  αναπαριστούν τα σήματα εξόδου, τα  $w_{i1}, w_{i2}, \dots, w_{nL}$  είναι οι συντελεστές βαρών που συνδέουν τον  $i$ -οστό κρυφό κόμβο με τους νευρώνες εισόδου και τα  $\beta_{i1}, \beta_{i2}, \dots, \beta_{np}$  οι

συντελεστές βαρών που συνδέουν τον  $i$ -οστό κρυφό κόμβο με τους νευρώνες εξόδου. Τέλος, τα  $b_1, b_2, \dots, b_L$  είναι τα κατώφλια των αντίστοιχων κρυφών κόμβων.

Σε όλες τις περιπτώσεις, για  $N$  αυθαίρετα, διαφορετικά δείγματα  $(x_i, t_i)$  όπου  $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbb{R}^n$  και  $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in \mathbb{R}^m$  με  $L$  κρυφούς κόμβους και συνάρτηση ενεργοποίησης  $g(x)$  το μαθηματικό μοντέλο εκφράζεται ως

$$\sum_{i=1..L} \beta_i \mathbf{G}(\mathbf{w}_i, \mathbf{b}_i, \mathbf{x}) = \mathbf{o}_j, j = 1, \dots, N \quad (2)$$

όπου  $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ ,  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ ,  $b_i$  το κατώφλι του  $i$ -οστού κρυφού κόμβου και  $G(w_i, b_i, x)$  η συνάρτηση ενεργοποίησης. Αξίζει να σημειωθεί ότι στον αλγόριθμο ELM τα βάρη εισόδου ( $w_i$ ) και τα κατώφλια (πολώσεις  $b_i$ ) του κρυφού στρώματος ανατίθενται **τυχαία** (1<sup>ο</sup> βήμα του αλγορίθμου), αν οι συναρτήσεις ενεργοποίησης είναι απείρως μη παραγωγίσιμες. Μετά την ανάθεση αυτή τα SLFNs μπορούν απλά να θεωρηθούν ένα **γραμμικό σύστημα** και άρα η (1) γράφεται ως εξής:

$$\sum_{i=1..L} \beta_i g(w_i^* x_j + b_i) = o_j, j = 1, \dots, N \quad (3)$$

, όπου  $G(w_i, b_i, x) = g(w_i^* x_j + b_i)$

Έχει αποδειχθεί ότι ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης με ένα κρυφό επίπεδο το οποίο έχει το πολύ  $L$  κρυφούς κόμβους (*hidden nodes*), με αυθαίρετη επιλογή των βαρών εισόδου και με σχεδόν οποιαδήποτε μη γραμμική συνάρτηση ενεργοποίησης μπορεί να εκπαιδευτεί από  $L$  ξεχωριστά παραδείγματα εκπαίδευσης με σχεδόν μηδενικό σφάλμα ( στην περίπτωση αυτή και σύμφωνα με την εικόνα 2 έχουμε  $N = L$ )

Συνεπώς, για  $L$  κρυφούς κόμβους η εκπαίδευση  $L$  δειγμάτων γίνεται με μηδενικό σφάλμα, δηλαδή

$$\sum_{j=1..L} \| o_j - t_j \| = 0 \quad (4)$$

Και συνεπώς υπάρχουν  $\beta_i, w_i, b_i$  τέτοια ώστε

$$\sum_{i=1..L} \beta_i g(w_i^* x_j + b_i) = t_j, j = 1, \dots, N \quad (5)$$

Η (3) μπορεί να γραφτεί και ως

$$H\beta = T, \quad (6)$$

όπου το  $H$  ονομάζεται μήτρα εξόδου του κρυφού επιπέδου του νευρωνικού δικτύου (*hidden layer output matrix*), είναι ορθογώνια και αντιστρέψιμη και η  $i$ -οστή στήλη της αναπαριστά το διάνυσμα εξόδου του  $i$ -οστού κρυφού νευρωνίου βάσει των εισόδων  $x_1, x_2, \dots, x_n$ . Πιο αναλυτικά:

$$H = \begin{pmatrix} g(w1 * x1 + b1) & \dots & g(wL * x1 + bL) \\ \vdots & \ddots & \vdots \\ g(w1 * xL + bL) & \dots & g(wL * xL + bL) \end{pmatrix} \quad (7)$$

Από την στιγμή που θα οριστούν τυχαία τα  $w, b$  (1<sup>ο</sup> βήμα του αλγορίθμου) είναι πολύ εύκολο να υπολογιστεί η μήτρα  $H$  (2<sup>ο</sup> βήμα του αλγορίθμου).

Στη συνέχεια, και με γνωστό πλέον το  $H$ , υπολογίζονται τα βάρη εξόδου  $\beta_i$  με βάση την εξίσωση (6) (που συνδέουν το κρυφό στρώμα με το στρώμα εξόδου των SLFNs) χρησιμοποιώντας μια απλή γενικευμένη αντιστροφή της μήτρας εξόδου  $H$  του κρυφού στρώματος (ορίζεται ως  $H^+$ )

$$\beta_i = H^+T \quad (8)$$

Οι μέθοδοι για τον υπολογισμό της μήτρας  $H^+$  (3<sup>ο</sup> βήμα του αλγορίθμου) ποικίλλουν και μπορεί να είναι η μέθοδος της ορθής προβολής, η μέθοδος ορθογωνιοποίησης, η επαναληπτική μέθοδος και η απλή αποσύνθεση αξίας. Όσον αφορά στην δεύτερη και την τρίτη μέθοδο συνήθως δεν χρησιμοποιούνται στον αλγόριθμο ELM, αφού η αναζήτηση και η επανάληψη προσπαθούν να αποφεύγονται. Η τέταρτη μέθοδος είναι αυτή που χρησιμοποιείται κατά κόρον.

Η μέθοδος απλής αποσύνθεσης αξίας, όπως και όλες οι υπόλοιπες μέθοδοι αποσύνθεσης (QR αποσύνθεση, Polar αποσύνθεση κ.α.) έχουν σαν στόχο το μετασχηματισμό των πινάκων σε πιο εύκολες και προσβάσιμες μορφές, διαφυλάττοντας συγχρόνως τις ιδιότητες τους, (όπως είναι η ορίζουσα, η τάξη ή η αντίστροφος), έτσι ώστε αυτές οι ποσότητες να υπολογίζονται αρκετά πιο εύκολα.

Η μήτρα  $H$  μπορεί να γραφτεί σύμφωνα με την παραπάνω μέθοδο με τη μορφή:

$$H = U\Sigma V^T \quad (9)$$

Όπου  $U$  είναι μία ορθογώνια μήτρα  $m \times m$ ,  $\Sigma$  μία μήτρα  $m \times n$  με μη αρνητικά στοιχεία στη διαγώνιο της και  $V^T$  η συζυγής μήτρα της  $V$ , που είναι επίσης ορθογώνια. Από τη στιγμή που θα υπολογιστούν οι μήτρες  $U$ ,  $V$ ,  $\Sigma$  με τη μέθοδο της ψευδοαντιστροφής, η οποία αναλύθηκε και τεκμηριώθηκε από τους Moore – Penrose, προκύπτει ότι ο ζητούμενος πίνακας  $H^+$  προκύπτει από την παρακάτω εξίσωση:

$$H^+ = V\Sigma^+U^T \quad (10)$$

Στις περισσότερες περιπτώσεις, τα δείγματα είναι πολύ περισσότερα από τον αριθμό των κρυφών κόμβων και η μήτρα  $H$  δεν είναι ορθογώνια και δεν υπάρχουν  $\beta_i$ ,  $w_i$ ,  $b_i$  τέτοια ώστε να ικανοποιείται η (4). Σε αυτές τις περιπτώσεις πρέπει να βρεθεί το ελάχιστο της νόρμας  $\|H\beta - T\|$  με άλλα λόγια το ελάχιστο  $\beta$  (4<sup>ο</sup> βήμα του αλγορίθμου)

$$\beta_{\min} = H^+T \quad (11)$$

Σύμφωνα με το θεώρημα των ελαχίστων νορμών για ένα γραμμικό σύστημα ισχύει το εξής:

$$\|x_0\| \leq \|x\|, \forall x \in \{x: \|Hx - y\| \leq \|Hz - y\|, \forall z \in \mathbb{R}^N\} \quad (12)$$



Συνεπώς για το  $\beta_{\min}$  θα ισχύει ότι θα έχει τη μικρότερη νόρμα από όλα τα υπόλοιπα  $\beta$  που αποτελούν λύση του συστήματος δηλαδή:

$$\|\beta_{\min}\| = \|H^+T\| \leq \|\beta\|, \forall \beta \in \{ \beta: \|H\beta - T\| \leq \|Hz - T\|, \forall z \in R^{N \times N} \} \quad (13)$$

Τα παραπάνω βήματα, δηλαδή:

- τυχαία επιλογή των τιμών των βαρών  $w_i$  και των πολώσεων  $b_i$
- υπολογισμός της μήτρας  $H$
- υπολογισμός της αντίστροφης μήτρας  $H^+$
- υπολογισμός του ελάχιστου  $\beta$

συνιστούν τον **αλγόριθμο της μηχανής μάθησης υψηλής απόδοσης (extreme learning machine ELM)**

Ο αλγόριθμος μπορεί να επεκταθεί πέρα από τα απλού επιπέδου νευρωνικά δίκτυα πρόσθιας τροφοδότησης (SLFN) και στα δίκτυα συναρτήσεων βάσης ακτινικού τύπου (Radial Basis Function - RBF). Στα SLFN επιλέγονται τυχαία τα βάρη  $w_i$  και οι πολώσεις  $b_i$ , ενώ στα RBF με την ίδια λογική επιλέγονται τυχαία τα κέντρα των πυρήνων και το εύρος  $\sigma$ . Αντίστοιχα το μαθηματικό μοντέλο εκφράζεται ως

$$\sum_{i=1..L} \beta_i \varphi_i(x_j) = o_j, j = 1, \dots, N \quad (14)$$

Όπου  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$  οι συντελεστές βαρών που συνδέουν τον  $i$ -οστό κρυφό κόμβο με τους νευρώνες εξόδου και  $\varphi(x)$  η έξοδος του  $i$ -οστού νευρωνίου που είναι συνήθως Gauss-ιανή

$$\varphi_i(x) = \varphi(\mu_i, \sigma_i, x) = \exp(-\|x - \mu_i\|^2 / \sigma_i) \quad (15)$$

με  $\mu_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{im}]^T$  το κέντρο του  $i$ -οστού πυρήνα και  $\sigma_i$  το εύρος του. Τα RBF δίκτυα έχουν την ίδια συμπεριφορά με τα SLFN όταν ο αριθμός των δειγμάτων είναι ίσος με τον αριθμό των κρυφών κόμβων συνεπώς ισχύει η (4) και άρα

$$\sum_{i=1..L} \beta_i \exp(-\|x - \mu_i\|^2 / \sigma_i) = t_j, j = 1, \dots, N \quad (16)$$

Καταλήγουμε δηλαδή και πάλι στην μορφή της (4) και το  $H$  ονομάζεται μήτρα εξόδου του κρυφού επιπέδου του RBF δικτύου.

Ο αλγόριθμος ELM ( για δίκτυα SLFN ή RBF ) δουλεύει αποδοτικά για οποιαδήποτε διαφορίσιμη ή μη , γραμμική ή μη συνάρτηση ενεργοποίησης. Για τους κρυφούς κόμβους των SLFN δικτύων, συνάρτηση ενεργοποίησης μπορεί να είναι οποιαδήποτε φραγμένη, μη σταθερή, τμηματικά συνεχής συνάρτηση (  $g:R \rightarrow R$  ) και για τους κρυφούς κόμβους των RBF δικτύων οποιαδήποτε ολοκληρώσιμη και τμηματικά συνεχής συνάρτηση (  $g:R \rightarrow R$  και  $\int g(x)dx \neq 0$  ) . Δεν απαιτείται για να είναι αποτελεσματικός ο αλγόριθμος η συνάρτηση να είναι συνεχής σε όλο το  $R$  αρκεί να είναι τμηματικά συνεχής (όπως οι συναρτήσεις κατωφλίου).

Επιπρόσθετα, σε αντίθεση με τις περισσότερες παραδοσιακές μεθόδους, αποδείχθηκε ότι οι παράμετροι σε κάθε δίκτυο είναι καλύτερο να επιλέγονται τυχαία από το να χάνεται πολύτιμος χρόνος προκειμένου να αποφασιστεί ποιες είναι οι αρχικές τιμές που θα πρέπει να πάρουν, καθώς και να υπάρχουν καθυστερήσεις λόγω του επαναυπολογισμού τους σε κάθε επανάληψη. Τέλος, είναι πολύ σημαντική η διαπίστωση ότι τα δίκτυα SLFN στα οποία προστίθενται τυχαία κρυφοί κόμβοι, λειτουργούν σαν καθολικοί προσεγγιστές (αλγόριθμος I-ELM). Αυτός είναι και ένας αποδοτικός τρόπος για την κατασκευή συνεχώς αυξανόμενων δικτύων πρόσθιας τροφοδότησης.

Όσον αφορά τα προβλήματα κατηγοριοποίησης (*classification problems*) και συγκρίνοντας τον αλγόριθμο ELM με τους με παραδοσιακούς ταξινομητές (*classifiers*) όπως για παράδειγμα τις μηχανές διανυσμάτων υποστήριξης (*Support Vector Machines SVM*), ή τον αλγόριθμο πίσω διάδοσης του λάθους (*Back Propagation BP*) παρατηρείται ότι είναι εκπληκτικά πιο γρήγορος, ξεπερνώντας το άνω φράγμα όσον αφορά στην ταχύτητα εκπαίδευσης των παραδοσιακών αλγορίθμων (όπως αναφέρθηκε παραπάνω). Εκτός από την επίτευξη του μικρότερου σφάλματος εκπαίδευσης, πράγμα το οποίο κάνουν και οι κλασικοί αλγόριθμοι για προβλήματα κατηγοριοποίησης, ο αλγόριθμος ELM έχει σαν στόχο και την επίτευξη της μικρότερης νόρμας των βαρών εισόδου καθώς έχει αποδειχθεί ότι όσο μικρότερη είναι η νόρμα αυτή, τόσο καλύτερη γενίκευση επιτυγχάνεται. Ένα άλλο κομμάτι στο οποίο υπερτερεί, είναι η χρησιμοποίηση και μη διαφορίσιμων συναρτήσεων ενεργοποίησης για τα SLFN χωρίς να χάνει καθόλου σε απόδοση. Η απόδοση βέβαια του αλγορίθμου ELM εξαρτάται από την επιλογή της συνάρτησης ενεργοποίησης (αν είναι σιγμοειδής, μη πολυωνυμική κτλ). Τέλος ο αλγόριθμος ELM ξεπερνά διάφορες δυσκολίες που συναντούν οι παραδοσιακοί αλγόριθμοι όπως η υπερπροσαρμοστικότητα και το τοπικό ελάχιστο. Όλα τα παραπάνω ισχύουν για τα νευρωνικά δίκτυα με ένα μόνο κρυφό επίπεδο κόμβων. Όταν υπάρχουν περισσότερα η απόδοση του ELM δεν είναι πολύ καλή και προτιμούνται οι παραδοσιακοί ταξινομητές. Ένα δεύτερο σημείο στο οποίο μειονεκτεί ο ELM είναι ότι δεν δίνει πολύ καλά αποτελέσματα όταν ο αριθμός των κρυφών κόμβων είναι μικρός.

## 4. Παραλλαγές του αλγορίθμου Extreme Learning Machine

Ο αλγόριθμος ELM επιτυγχάνει πολύ καλούς χρόνους εκπαίδευσης και καλύτερη απόδοση γενίκευσης, όπως προαναφέρθηκε παρ' όλα αυτά υπάρχουν πολλά περιθώρια βελτίωσης. Μεταγενέστερα αναπτύχθηκαν αρκετές παραλλαγές του ELM οι οποίες μπορούν να συνδυάζουν και άλλους αλγόριθμους μαζί για καλύτερα αποτελέσματα, η καθεμιά αναφέρεται σε συγκεκριμένη κατηγορία προβλημάτων (οποιαδήποτε προβλήματα και εφαρμογές αναφέρονται παρακάτω εμπίπτουν στην κατηγορία προβλημάτων κατηγοριοποίησης), και ασχολείται με ένα διαφορετικό κομμάτι του αλγορίθμου (για παράδειγμα με τη συνάρτηση ενεργοποίησης, με τον αριθμό των κρυφών κόμβων κτλ)

### 4.1 E-ELM (Evolutionary ELM)

Για να μπορέσει να αντιμετωπιστεί η προβληματική απόδοση του ELM όταν ο αριθμός των κρυφών κόμβων είναι μικρός, ένας υβριδικός αλγόριθμος που έχει αναπτυχθεί είναι ένας συνδυασμός ELM και αλγορίθμου διαφορικής εξέλιξης (differential evolution DE). Ο αλγόριθμος DE χρησιμοποιείται για να βρει τις βέλτιστες τιμές των βαρών εισόδου και των πολώσεων των κρυφών κόμβων και ο αλγόριθμος ELM για τον υπολογισμό των βαρών εξόδου.

Ο αλγόριθμος διαφορικής εξέλιξης (DE) είναι ένας στοχαστικός αλγόριθμος ολικής βελτιστοποίησης βασισμένος στην εξέλιξη ενός πληθυσμού και έχει τρία βασικά πλεονεκτήματα: βρίσκει το πραγματικό ελάχιστο ανεξαρτήτως των αρχικών τιμών των παραμέτρων, παρουσιάζει γρήγορη σύγκλιση και κάνει χρήση λίγων παραμέτρων για τον έλεγχό του. Η απλότητά του, η ταχύτητα και η ευκολία στην χρήση του είναι κάποια επιπλέον χαρακτηριστικά του.

Ένας πληθυσμός στον DE αποτελείται από NP διανύσματα

$$\{\theta_{i,G} | i=1,2,\dots, NP\} \quad (17)$$

όπου G είναι ο αριθμός των γενεών. Ο πληθυσμός αρχικοποιείται τυχαία από μια ομοιόμορφη κατανομή και εξελίσσεται σε κάθε γενιά με τη χρήση τριών τελεστών: μετάλλαξη (mutation), διασταύρωση (crossover) και επιλογή (selection).

Στη μετάλλαξη το διάνυσμα προκύπτει από την παρακάτω ισότητα:

$$v_{i,G+1} = \theta_{r_1,G} + F * (\theta_{r_2,G} - \theta_{r_3,G}) \quad (18)$$

Στη διασταύρωση το διάνυσμα προκύπτει από την παρακάτω ισότητα:

$$\mu_{i, G+1} = \begin{cases} \nu_{j_i, G+1}, & \text{αν } \text{randb}(j) \leq CR \quad (19) \\ \theta_{j_i, G}, & \text{αν } \text{randb}(j) > CR \quad (20) \end{cases}$$

Τέλος στην περίπτωση της επιλογής αν το διάνυσμα  $\mu_{i, G+1}$  είναι καλύτερο από το  $\theta_{i, G}$ , τότε το  $\theta_{i, G+1}$  παίρνει τη θέση του  $\mu_{i, G+1}$ . Σε αντίθετη περίπτωση γίνεται το  $\theta_{i, G+1}$  παίρνει την τιμή του  $\theta_{i, G}$ .

Όσον αφορά στη υβριδική προσέγγιση του μοντέλου της E-ELM, αρχικά δημιουργείται τυχαία ο πληθυσμός. Κάθε μονάδα του πληθυσμού, αποτελείται από ένα σύνολο βαρών  $w$  και πολώσεων  $b$ . Στη συνέχεια τα βάρη εξόδου υπολογίζονται με βάση τον αλγόριθμο ELM. Αφού υπολογιστούν τα βάρη εξόδου, υπολογίζεται η καταλληλότητα (fitness) του κάθε στοιχείου του πληθυσμού, η οποία σχετίζεται άμεσα με τη συνάρτηση κόστους  $E$  (ρίζα του μέσου τετραγωνικού σφάλματος). Μόλις υπολογιστούν οι καταλληλότητες όλων των μονάδων του πληθυσμού εφαρμόζονται τα τρία στάδια του αλγόριθμου διαφορικής εξέλιξης που αναφέρθηκαν παραπάνω. Επειδή το κριτήριο υπολογισμού καταλληλότητας είναι μόνο ένα (σχετίζεται με τη συνάρτηση κόστους), δεν είναι αρκετό για να επιτευχθεί καλύτερη απόδοση γενίκευσης και για το λόγο αυτό προστίθεται και το κριτήριο της μικρότερης νόρμας των βαρών εξόδου.

Συγκρινόμενος με τον κλασικό αλγόριθμο ELM, ο αλγόριθμος E-ELM προσφέρει μία πιο συμπαγή αρχιτεκτονική δικτύου, πράγμα που εξασφαλίζει μεγαλύτερη ταχύτητα απόκρισης το οποίο είναι βοηθητικό σε προβλήματα ταξινόμησης, όπου τα δεδομένα είναι πάρα πολλά, αλλά και το μέγεθος του συνόλου εκπαίδευσης άγνωστο. Πέρα από αυτό εξασφαλίζεται καλύτερη απόδοση γενίκευσης.

## 4.2 C-ELM (Complex ELM)

Στα υψηλής ταχύτητας ψηφιακά επικοινωνιακά συστήματα, οι ισοσταθμιστές χρησιμοποιούνται πολύ συχνά στους δέκτες για να ανακτηθούν οι αυθεντικές συμβολοσειρές από τα σήματα λήψης. Τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης (SLFN), τα δίκτυα συναρτήσεων βάσης ακτινικού τύπου (RBF) και τα επαναλαμβανόμενα νευρωνικά δίκτυα (*recurrent neural networks*) χρησιμοποιούνται σε μεγάλο βαθμό για την επίλυση προβλημάτων σχετικών με τους ισοσταθμιστές αφού είναι πολύ αποδοτικά σε μη γραμμικά προβλήματα ταξινόμησης.

Ο αλγόριθμος C-ELM επεκτείνει τον αλγόριθμο ELM από τον απλό τομέα (*real domain*) σε πολύπλοκο τομέα (*complex domain*) και χρησιμοποιεί πολύπλοκες συναρτήσεις ενεργοποίησης. Τα βάρη εισόδων και οι πολώσεις κρυφών κόμβων επιλέγονται τυχαία

και τα βάρη εξόδου υπολογίζονται όπως και στον ELM. Χρησιμοποιείται κατά κόρον για την ισοστάθμιση πολύπλοκων μη γραμμικών καναλιών για τη διαμόρφωση πλάτους τετραγωνικών σημάτων (*Quadrature Amplitude Modulation QAM*). Ο αλγόριθμος C-ELM ολοκληρώνει τη φάση της μάθησης εξαιρετικά γρήγορα και με πολύ μικρότερο σφάλμα συγκρινόμενος με άλλους αντίστοιχους ισοσταθμιστές όπως ο CMRAN (*Complex Minimal Resource Allocation Network*), ο CRBF (*Complex Radial Basis Function*) και ο CBP (*Complex Back Propagation*). Ο τελευταίος είναι αρκετά ευαίσθητος στο μέγεθος του ρυθμού εκπαίδευσης και στο μέγεθος των βαρών εισόδου, παράμετροι που αφήνουν σχεδόν ανεπηρέαστο τον αλγόριθμο C-ELM ο οποίος επιπρόσθετα αποφεύγει το εμπόδιο του τοπικού ελάχιστου, εφαρμόζεται εύκολα και γρήγορα έχοντας υψηλούς χρόνους απόκρισης παράγοντας πολύ βασικός για την καλή απόδοση ενός ισοσταθμιστή. Οι αποδόσεις του C-ELM ποικίλλουν ανάλογα με την πολύπλοκη συνάρτηση ενεργοποίησης που θα επιλεγεί.

#### **4.3 OS - ELM / EOS-ELM / OS – Fuzzy – ELM (Online Sequential / Enhanced Online Sequential / On Line Sequential Fuzzy ELM)**

Σε πολλές περιπτώσεις, εφαρμογές των SLFN δικτύων χρησιμοποιούν τη μέθοδο της μαζικής μάθησης (*batch learning*), η οποία χαρακτηρίζεται χρονοβόρα, καθώς απαιτεί αρκετές επαναλήψεις στο σύνολο των δεδομένων εκπαίδευσης. Αρκετός χρόνος απαιτείται και στην επιλογή των κατάλληλων παραμέτρων και κριτηρίων ώστε να υπάρξει η επιθυμητή σύγκλιση σε σύντομο χρονικό διάστημα. Κάθε φορά που φτάνει ένα νέο πακέτο δεδομένων οι παραδοσιακοί αλγόριθμοι μάθησης χρησιμοποιούν τα προηγούμενα δεδομένα μαζί με τα νέα και εκτελούν μία επανεκπαίδευση. Για τους παραπάνω λόγους από πολλές εταιρείες προτιμούνται online ακολουθιακοί αλγόριθμοι εκμάθησης συγκριτικά με τους αλγόριθμους μαζικής μάθησης (χαρακτηριστικό παράδειγμα των τελευταίων αποτελεί ο BP και ο στοχαστικός BP (SGBP)).

Ένας online ακολουθιακός αλγόριθμος που αποτελεί παραλλαγή του κλασικού αλγορίθμου ELM, είναι ο OS-ELM ο οποίος μπορεί να εφαρμοστεί και σε κρυφούς κόμβους SLFN δικτύων και σε κρυφούς κόμβους RBF δικτύων χρησιμοποιώντας διαφορίσιμες συναρτήσεις ενεργοποίησης και μη. Ο OS-ELM μπορεί να εκπαιδεύει με τα πακέτα δεδομένων να έρχονται όχι μόνο ένα προς ένα (*one by one*) αλλά και κομμάτι- κομμάτι (*chunk by chunk*) με σταθερό ή και μεταβλητό μέγεθος κομματιού. Τα παραδείγματα εκπαίδευσης είναι διαδοχικά και κάθε χρονική στιγμή μόνο το νέο «κομμάτι» δεδομένων εκπαιδεύει το δίκτυο χωρίς να λαμβάνεται υπόψη τίποτα άλλο. Κάθε πακέτο ή κομμάτι αφού χρησιμοποιηθεί για την εκπαίδευση πετιέται και δεν επαναχρησιμοποιείται. Τέλος, ο συγκεκριμένος αλγόριθμος εκμάθησης δεν έχει καμία πληροφορία για το πόσα θα είναι τα δεδομένα εκπαίδευσης. Το μόνο που πρέπει να οριστεί εξ αρχής είναι ο αριθμός των κρυφών κόμβων και να θεωρηθεί ότι όλα τα δεδομένα είναι διαθέσιμα για εκπαίδευση.

Συγκρινόμενος με τους υπόλοιπους ακολουθιακούς αλγόριθμους πάνω σε προβλήματα ταξινόμησης (SGBP, RAN, RANЕКF, MRAN) ο OS-ELM επιτυγχάνει καλύτερη γενίκευση και έχει πολύ μεγαλύτερη ταχύτητα εκπαίδευσης. Στην εκπαίδευση one - by - one ο SGBP έχει παρόμοιους χρόνους, αλλά στην εκπαίδευση chunk by chunk ο OS-ELM υπερτερεί ξεκάθαρα και μάλιστα χωρίς να χρειάζεται απαραίτητα σταθερό μήκος κομματιού. Έχει παρατηρηθεί ότι οι χρόνοι του αλγόριθμου βελτιώνονται κατά πολύ όταν αυξάνεται το μήκος του κομματιού. Οι χρόνοι του συγκριτικά με τον κλασικό αλγόριθμο ELM είναι σχετικά κοντά.

Οι χρόνοι του OS-ELM διαφέρουν και δεν είναι πάντα πολύ καλοί, εξαρτάται κάθε φορά από το πείραμα και την προσομοίωση. Για να μπορέσει να επιτευχθεί μία πιο σταθερή και αξιόπιστη συμπεριφορά αναπτύχθηκε ο αλγόριθμος EOS-ELM, ο οποίος συνδυάζει πολλά OS-ELM δίκτυα, τα οποία έχουν τον ίδιο αριθμό κρυφών κόμβων και την ίδια συνάρτηση ενεργοποίησης για κάθε κρυφό κόμβο. Το κάθε δίκτυο OS εκπαιδεύεται με νέα δεδομένα, τα οποία καταφθάνουν ακολουθιακά και φυσικά οι παράμετροι του δικτύου επιλέγονται τυχαία. Σε ένα δίκτυο OS-ELM η μόνη παράμετρος που πρέπει να τεθεί από την αρχή είναι ο αριθμός των κρυφών κόμβων. Αντίστοιχα στο δίκτυο EOS-ELM ορίζεται αρχικά μόνο ο βέλτιστος αριθμός κρυφών κόμβων, ο οποίος είναι κοινός για όλα τα επιμέρους OS. Η έξοδος του συνολικού συνδυασμού δικτύων είναι ουσιαστικά ο μέσος όρος των εξόδων κάθε δικτύου OS.

Είναι αναμενόμενο ένα δίκτυο EOS, να αποδίδει καλύτερα από ότι ένα μεμονωμένο δίκτυο OS. Το κάθε δίκτυο OS έχει διαφορετική προσαρμοστικότητα στο πακέτο δεδομένων που έρχεται για εκπαίδευση. Κάποια προσαρμόζονται πιο γρήγορα, κάποια όχι και τόσο, βλέποντας τα όμως συνολικά, αυτά με την όχι τόσο καλή απόδοση «απορροφώνται» από τα «καλά». Σύμφωνα με τη στατιστική, ένας πληθυσμός είναι πιο κοντά στο επιθυμητό αποτέλεσμα από ότι τα μεμονωμένα στοιχεία. Συνεπώς συμπεραίνεται ότι τα αποτελέσματα του EOS-ELM δικτύου έχουν μικρότερο εύρος διακύμανσης και άρα είναι πιο κοντά στο βέλτιστο αποτέλεσμα. Όταν ο αριθμός των δικτύων που συμμετέχουν στο συνολικό αυξάνεται, τότε το αποτέλεσμα είναι ακόμα πιο κοντά στο επιθυμητό. Από την άλλη πλευρά, όσο πιο μεγάλος είναι αυτός ο αριθμός τόσο πιο πολύπλοκο γίνεται το δίκτυο και τόσο πιο αργοί οι υπολογισμοί.

Σε γενικές γραμμές, με αυτόν τον τρόπο αντιμετώπισης του προβλήματος επιτυγχάνεται καλύτερη απόδοση γενίκευσης, μικρότεροι χρόνοι εκπαίδευσης και καλύτερη συνεκτικότητα, ακρίβεια και σταθερότητα κάτω από οποιεσδήποτε συνθήκες.

Τα ασαφή συστήματα συμπερασμού (*Fuzzy Inference Systems FIS*) χρησιμοποιούνται ολοένα και περισσότερο για την προσέγγιση συναρτήσεων σε προβλήματα παλινδρόμησης (*regression*) και ταξινόμησης (*classification*). Οι αλγόριθμοι που εφαρμόζονται και εδώ, είναι είτε αλγόριθμοι μαζικής μάθησης (*batch learning*), είτε online ακολουθιακοί αλγόριθμοι. Συνδυάζοντας τα πλεονεκτήματα ενός νευρωνικού δικτύου (έχουν σχέση με την εκπαίδευση), και τα πλεονεκτήματα ενός ασαφούς

συστήματος (έχουν σχέση με την προσέγγιση συναρτήσεων), μπορεί να δημιουργηθεί ένα υβριδικό μοντέλο, το οποίο θα εκμεταλλεύεται τα πλεονεκτήματα και των δύο. Με βάση αυτό το νεύρο-ασαφές υβριδικό δίκτυο αναπτύχθηκε ο αλγόριθμος OS-Fuzzy-ELM, ο οποίος ουσιαστικά αποτελεί επέκταση του OS-ELM.

Ο αλγόριθμος OS-Fuzzy-ELM αποτελείται ουσιαστικά από δύο φάσεις. Η πρώτη φάση ονομάζεται αρχικοποίηση (*initialization phase*), στην οποία το ασαφές δίκτυο εκπαιδεύεται βάση του ασαφούς ELM αλγόριθμου με μαζική άφιξη δεδομένων τα οποία δεν χρησιμοποιούνται ξανά αφού τελειώσει η εκπαίδευση. Ο απαιτούμενος αριθμός δεδομένων εκπαίδευσης, στο συγκεκριμένο στάδιο, πρέπει να είναι πολύ μικρός για να είναι ίσος με τον αριθμό κανόνων του συστήματος. Στο δεύτερο στάδιο ακολουθεί η εκπαίδευση από τον αλγόριθμο OS-Fuzzy-ELM ένα προς ένα ή κομμάτι-κομμάτι και τα δεδομένα δεν επαναχρησιμοποιούνται.

Συγκρινόμενος με άλλους αλγόριθμους, σε προβλήματα ταξινόμησης όπως τον ANFIS, τον Simpl\_eTS και τον e\_TS παρατηρείται ότι ο OS-Fuzzy-ELM απαιτεί πάντα αρκετά μικρότερο αριθμό κανόνων από τους υπόλοιπους, η ακρίβεια της εκπαίδευσης είναι παρόμοια και στους 4 αλγόριθμους ενώ στην ταχύτητα υπερτερεί εμφανώς ο OS-Fuzzy-ELM.

#### 4.4 SCELM (Sine / Cosine ELM)

Ένας αλγόριθμος ο οποίος κωδικοποιεί και ενσωματώνει πληροφορίες εξαρχής και με αυτόν τον τρόπο επεκτείνει τον κλασικό ELM, είναι ο CELM. Οι συναρτήσεις ενεργοποίησης των κρυφών κόμβων εδώ είναι συναρτήσεις ημιτόνου και συνημίτονου βασιζόμενες στο θεώρημα επέκτασης σειρών Fourier. Η συνάρτηση εξόδου δίδεται παρακάτω:

$$\varphi(x) = \sum_{k=1 \dots n} (w_{02k} \cos(k(w_{h2k} x)) + w_{02k-1} \sin(k(w_{h2k-1} x))) + w_{02n+1} \quad (21)$$

όπου  $w_{0i}$  είναι το βάρος του  $i$ -οστού νευρωνίου που αντιστοιχεί από τον κρυφό κόμβο στην έξοδο, και  $w_{hi}$  είναι το βάρος του  $i$ -οστού νευρωνίου από την είσοδο στον κρυφό κόμβο. Η έξοδος είναι γραμμική.

Ο αλγόριθμος SCELM έχει δύο διαφορετικές εκφάνσεις. Όταν η συνάρτηση ενεργοποίησης είναι άρτια, η συνάρτηση μεταφοράς του  $k$ -οστού κρυφού νευρωνίου είναι  $\cos kx$  και του οποιουδήποτε  $(n+1)$  κρυφού νευρωνίου είναι 1. Ο αλγόριθμος αυτός

ονομάζεται Cosine ELM (CELM). Αντίστοιχα όταν η συνάρτηση ενεργοποίησης είναι περιπτή, η συνάρτηση μεταφοράς του k-οστού κρυφού νευρωνίου είναι  $\sin kx$  και του οποιουδήποτε (n+1) κρυφού νευρωνίου είναι 1. Ο αλγόριθμος τότε ονομάζεται Sine ELM (SELM).

Η συγκεκριμένη παραλλαγή έχει καλύτερη απόδοση γενίκευσης, ταχύτερους χρόνους εκπαίδευσης και ρυθμό σύγκλισης από την κλασική ELM. Θεωρείται δεδομένο ότι συγκριτικά με τους παραδοσιακούς αλγόριθμους σύγκλισης είναι εκατοντάδες φορές ταχύτερη.

#### **4.5 I-ELM/CI-ELM / EI-ELM (Incremental / Convex Incremental /Enhanced Incremental ELM)**

Όπως αναφέρθηκε στο πρώτο κεφάλαιο, μία παραλλαγή του κλασικού ELM, είναι ο I-ELM ο οποίος ουσιαστικά προσθέτει με τυχαίο τρόπο κρυφά νευρώνια στο δίκτυο και στην συνέχεια υπολογίζει τις εξόδους, χωρίς να επαναυπολογίζει τα βάρη εξόδου όλων των κόμβων από την στιγμή που προστέθηκε το κρυφό νευρώνιο. Οι κρυφές παράμετροι δεν είναι ανεξάρτητες μόνο με το σύνολο των δεδομένων εκπαίδευσης, αλλά και μεταξύ τους. Από την στιγμή που θα οριστούν τα κέντρα  $a_i$  και τα εύρη  $b_i$  υπολογίζεται το βάρος εξόδου  $\beta_i$  το οποίο δεν ξανά αλλάζει. Οι συναρτήσεις ενεργοποίησης που χρησιμοποιεί μπορεί να είναι διαφορίσιμες ή μη, συνεχείς ή τμηματικά συνεχείς.

Το μαθηματικό μοντέλο της συνάρτησης εξόδου είναι

$$f_n(\mathbf{x}) = f_{n-1}(\mathbf{x}) + \beta_n * g_i(\mathbf{x}) \quad (22)$$

Ο αλγόριθμος CI-ELM αντίστοιχα με τον I-ELM προσθέτει με τυχαίο τρόπο κρυφούς κόμβους στο δίκτυο και υπολογίζει τα βάρη εξόδου, ενώ και εδώ όλες οι υπόλοιπες παράμετροι επιλέγονται τυχαία. Η βασική διαφορά με τον I-ELM, είναι ότι αφού προστεθεί ένας καινούριος κόμβος τα βάρη όλων των προηγούμενων επαναυπολογίζονται. Με τη μέθοδο αυτή έχουμε ταχύτερη σύγκλιση (και για τα SLFN και για τα RBF δίκτυα) διατηρώντας την απλότητα και τη συνοχή του νευρωνικού δικτύου. Η ιδιότητα του καθολικού προσεγγιστή δεν παύει να ισχύει και εδώ.

Μία τελευταία παραλλαγή που μοιάζει πολύ με τον I-ELM είναι ο EI-ELM, ο οποίος βρίσκει εφαρμογή και σε μη νευρωνικά δίκτυα. Ο I-ELM μπορεί να θεωρηθεί υποπερίπτωση του EI-ELM για  $k = 1$ . Ουσιαστικά και ο EI-ELM προσθέτει με τυχαίο τρόπο κρυφούς κόμβους στο δίκτυο, όμως σε κάθε στάδιο μάθησης, διαλέγει τον βέλτιστο κρυφό κόμβο που θα χρησιμοποιήσει προκειμένου να ελαχιστοποιήσει το



σφάλμα. Επιτυγχάνει καλύτερη ταχύτητα σύγκλισης και δημιουργία πιο συνεκτικών και συνεπών δικτύων. Τέλος μπορεί να συνδυαστεί με τον CI-ELM.

#### 4.6 OP-ELM / P-ELM (Optimally Pruned / Pruned ELM)

Ο όγκος των δεδομένων αυξάνεται ολοένα και περισσότερο σε όλα τα επιστημονικά πεδία. Η αποθήκευση των δεδομένων, η επεξεργασία τους για την εξόρυξη κανόνων και άλλων συμπερασμάτων καθώς και οι διάφοροι υπολογισμοί γίνονται ολοένα και πιο δύσκολοι. Τα απλά νευρωνικά δίκτυα πρόσθιας τροφοδότησης (*FNN*) πολλές φορές χαρακτηρίζονται «αργά» για όλες αυτές τις απαιτήσεις και τα μη γραμμικά μοντέλα, παρά την καλή τους απόδοση, δεν μπορούν να χρησιμοποιηθούν ευρέως για την αντιμετώπιση των παραπάνω προβλημάτων.

Ο αλγόριθμος OP-ELM έχει σαν βάση του τον ELM, άλλα διαφέρει ως προς αυτόν στις συναρτήσεις ενεργοποίησης που χρησιμοποιεί (εκτός από σιγμοειδείς που χρησιμοποιεί ο κλασικός ELM, Gauss-ιανές, γραμμικές). Η συνάρτηση εξόδου είναι γραμμική. Οι παράμετροι των κρυφών κόμβων επιλέγονται και εδώ τυχαία και φροντίζεται κατάλληλα ώστε ο αρχικός αριθμός των κρυφών νευρώνων που θα επιλεγεί, να είναι ξεκάθαρα μεγαλύτερος από τον αριθμό των μεταβλητών του συστήματος (αφού στο επόμενο βήμα τα άχρηστα κρυφά νευρώνια θα αποκοπούν).

Για την αποκοπή των άχρηστων κρυφών νευρώνων χρησιμοποιείται ο αλγόριθμος πολλαπλών αποκρίσεων αραιής παλινδρόμησης (*Multiresponse Sparse Regression MRSR*). Πιο συγκεκριμένα χαρακτηρίζεται ως αλγόριθμος κατάταξης μεταβλητών, πάρα επιλογής, και είναι επέκταση του αλγορίθμου παλινδρόμησης ελάχιστης γωνίας (*Least Angle Regression*). Βασικό χαρακτηριστικό του αλγορίθμου MRSR είναι ότι δίνει λύσεις σε γραμμικά προβλήματα.

Στην συνέχεια αφού η κατάταξη των κρυφών νευρώνων είναι έτοιμη, η απόφαση για το ποια θα επιλεγθούν να συμμετέχουν ενεργά στο νευρωνικό δίκτυο βασίζεται στον αλγόριθμο Leave One - Out (LOO). Το πρόβλημα του συγκεκριμένου είναι πως καταναλώνει πολύ χρόνο ώσπου να δώσει ένα αποτέλεσμα σε μεγάλα σύνολα δεδομένων, πράγμα που διορθώνεται με τη λεγόμενη Πρόβλεψη Αθροίσματος Τετραγώνων (*PREdiction Sum of Squares*)

Ο αλγόριθμος OP-ELM, συγκρινόμενος με αντίστοιχους αλγόριθμους, τον LS-SVM (*Least - Squares Support Vector Machine*) και τον MLP (*Multilayer Perceptron*) σε προβλήματα ταξινόμησης και κατηγοριοποίησης, έχει σχεδόν ίδιο επίπεδο ακρίβειας αλλά υπερέχει πολύ στον χρόνο που καταναλώνει για να φέρει τα αποτελέσματα και εκεί η διαφορά είναι τεράστια.

Μία παραλλαγή η οποία βασίζεται επίσης στην αποκοπή κάποιων κρυφών νευρώνων οι οποίοι φαίνονται να μην επηρεάζουν καθόλου, ή να επηρεάζουν λίγο την έξοδο του δικτύου είναι ο P-ELM (*Pruned Extreme Learning Machine*). Βασίζεται σε στατιστικές

μεθόδους και κριτήρια και επιτυγχάνει ελάχιστα αισθητά καλύτερα αποτελέσματα σε προβλήματα ταξινόμησης.

#### 4.7 R-ELM (Regularized ELM)

Όλες οι παραλλαγές του ELM, που έχουν προαναφερθεί έχουν σαν στόχο την βελτίωση του κλασικού αλγορίθμου όσον αφορά στην ακρίβεια, τους χρόνους εκπαίδευσης και συνολικά την καλύτερη του απόδοση σε προβλήματα ταξινόμησης. Ο κλασικός αλγόριθμος παρουσιάζει κάποιες αδυναμίες που δημιουργούν προβλήματα στην πορεία. Το πρώτο είναι ότι βασίζεται αποκλειστικά και μόνο στον προσδιορισμό και ελαχιστοποίηση του εμπειρικού ρίσκου (*empirical risk minimization*) για να μπορέσει να ξεπεράσει το πρόβλημα της υπερπροσαρμοστικότητας (*overfitting*). Το δεύτερο είναι ότι ο ELM παρουσιάζει περιορισμένες δυνατότητες ελέγχου αφού υπολογίζει απευθείας την ελάχιστη νόρμα βάση της μεθόδου των ελαχίστων τετραγώνων. Τέλος μπορεί να οδηγήσει σε αποτελέσματα λιγότερο αξιόπιστα λόγω φαινομένων ετεροσκεδαστικότητας ή λόγω ύπαρξης ακραίων εσφαλμένων τιμών (*outliers*).

Σύμφωνα με τη στατιστική θεωρία μάθησης, η αληθινή πρόβλεψη ρίσκου μάθησης ορίζεται σαν το άθροισμα του εμπειρικού ρίσκου και του κατασκευαστικού ρίσκου. Ένα καλό μοντέλο θα πρέπει να βρίσκει την ισορροπία μεταξύ των δύο αυτών ρίσκων. Το ποσοστό παρουσίας του καθενός μπορεί να οριστεί από έναν παράγοντα  $\gamma$ , ο οποίος εμφανίζεται και στον τύπο του εμπειρικού και του κατασκευαστικού ρίσκου. Αυτό ακριβώς υπαγορεύει και ο αλγόριθμος R-ELM. Συνεπώς

$$\beta = (I/\gamma + HTD^2H)T^{-1} HT^{-1} D^2 T^{-1} (23)$$

$$\epsilon_i = \alpha_i / \gamma (24), \text{ όπου}$$

$$\alpha_i = -\gamma(H\beta - T)T^{-1} (25)$$

Ο αλγόριθμος R-ELM βασισμένος σε αυτή την λογική δεν διατηρεί απλά το πλεονέκτημα της ταχύτατης εκπαίδευσης, αλλά επιτυγχάνει καλύτερη απόδοση γενίκευσης από ότι ο ELM. Η ικανότητα ελέγχου γίνεται ακόμα μεγαλύτερη λόγω της παρουσίας του παράγοντα  $\gamma$ , όπως προαναφέρθηκε και ο θόρυβος που προκαλείται από ακραίες εσφαλμένες τιμές στο πακέτο των δεδομένων προς εκπαίδευση εξαλείφεται. Τέλος η λογική του αλγορίθμου R-ELM μπορεί να συνδυαστεί και με τον αλγόριθμο I-ELM.

#### 4.8 SB-ELM (Sparse-Bayesian ELM)

Το μειονέκτημα του κλασικού αλγορίθμου ELM και όσων παραλλαγών αναφέρθηκαν παραπάνω για το οποίο δεν έχει παρουσιαστεί ακόμα λύση είναι το ότι το αποτέλεσμα

επηρεάζεται άμεσα από τον αριθμό των κρυφών κόμβων (όσο μικρότερος είναι ο αριθμός τόσο χειρότερα τα αποτελέσματα του αλγόριθμου). Το πρόβλημα αυτό επιλύει η παραλλαγή Sparse Bayesian ELM (SBELM), η οποία δεν προσθαφαιρεί κόμβους στο δίκτυο αλλά προσπαθεί να προβλέψει με στατιστική και πιθανότητες ποια θα είναι η έξοδος του δικτύου. Ουσιαστικά ο υπολογισμός του  $W$  γίνεται με την κατά Bayes προσέγγιση και όχι με υπολογισμό της μήτρας  $H$  όπως στον κλασικό αλγόριθμο. Στόχος της είναι να αυξήσει την πιθανότητα της εξόδου στο μέγιστο και με αυτόν τον τρόπο νικά την υπερπροσαρμοστικότητα και επιτυγχάνει μεγαλύτερη ακρίβεια.

#### 4.9 S-ELM (Sparse ELM)

Η λύση που δίνει στα προβλήματα ταξινόμησης ο αλγόριθμος ELM είναι περίπλοκη και συνεπώς χρειάζεται μεγάλος αποθηκευτικός χώρος και αρκετός χρόνος εκπαίδευσης για μεγάλοι όγκου δεδομένων εφαρμογές. Επιπρόσθετα, ο κλασικός ELM βασίζει τη λύση που δίνει στην αντιστροφή της μήτρας  $H$  όπου η πολυπλοκότητα της κυμαίνεται από την τετραγωνική στην κυβική, ανάλογα με το μέγεθος του συνόλου δεδομένων εκπαίδευσης (παρά όλα αυτά παραμένει πιο αποδοτικός από τους παραδοσιακούς αλγόριθμους).

Ο εναλλακτικός αλγόριθμος Sparse ELM, ουσιαστικά ενοποιεί διαφορετικές θεωρίες ταξινόμησης και εφαρμόζεται και σε δίκτυα SLFN και σε δίκτυα RBF, ελαχιστοποιεί τον αποθηκευτικό χώρο που χρειάζεται και δημιουργεί ένα πιο συνεπές δίκτυο. Αρχικά πραγματοποιείται μία αντιστοίχιση χαρακτηριστικών από το σύνολο εισόδου σε ένα περισσότερων διαστάσεων σύνολο. Η συγκεκριμένη συνάρτηση αντιστοίχισης  $h(x)$ , μπορεί να δημιουργηθεί τυχαία και να μην είναι γνωστή εξαρχής. Όπως έχει αποδειχθεί σε οποιαδήποτε σύνολα στο χώρο  $R^d$  που η τομή τους είναι μηδενική μπορεί να υπάρχει συνάρτηση συνεχής,  $f(x)$ , η οποία μπορεί να τα χωρίσει. Ο αριθμός των κρυφών κόμβων ορίζεται από την αρχή και δεν αλλάζει. Τα σφάλματα εκπαίδευσης  $\xi$  επιτρέπονται στην Sparse ELM. Η υπερπροσαρμογή ή αλλιώς υπερεκπαίδευση, αντιμετωπίζεται ελαχιστοποιώντας το άθροισμα των εμπειρικών σφαλμάτων και των κατασκευαστικών ρίσκων.

Συγκρίνοντας τον Sparse ELM, με τους κλασικούς ταξινομητές, δηλαδή τις μηχανές διανυσμάτων υποστήριξης (SVM), ο πρώτος σε αντίθεση με τους δεύτερους δεν έχει τον περιορισμό του αθροίσματος  $\sum_{i=1..N} \alpha_i t_i = 0$  και συνεπώς ψάχνει για τη βέλτιστη λύση σε ένα μεγαλύτερο εύρος τιμών. Είναι ένα κλασικό πρόβλημα τετραγωνικού προγραμματισμού. Στην παραλλαγή αυτή του ELM, το συνολικό πρόβλημα χωρίζεται σε υποπρόβλήματα και στο κάθε υποπρόβλημα μεταβάλλεται μία μεταβλητή  $\alpha_i$  ώσπου το άθροισμα του παραπάνω περιορισμού εξαλειφθεί. Επίσης ο υπολογισμός του  $\beta$  εδώ, είναι αρκετά πιο απλός και συγκριτικά με τους SVM's επιτυγχάνει μεγαλύτερη απόδοση γενίκευσης, καλύτερους χρόνους εκπαίδευσης, όχι όμως και καλύτερη απόδοση γενικά.

Η παραλλαγή του αλγορίθμου (Sparse ELM) συγκρινόμενη με τον κλασικό αλγόριθμο ELM (σε προβλήματα ταξινόμησης πάντα), επιτυγχάνει παρόμοια απόδοση γενίκευσης.

Ο κλασικός ELM λύνει τα προβλήματα απευθείας ενώ ο Sparse ELM χρησιμοποιεί επαναληπτικές μεθόδους συνδυάζοντας και τον αλγόριθμο OAO πράγμα που δεν τον κάνει τόσο ανταγωνιστικό σε θέματα ταξινόμησης, ούτε όσον αφορά το χρόνο εκπαίδευσης ούτε όσον αφορά το χρόνο ελέγχου. Συνεπώς δεν αποτελεί καλή επιλογή για προβλήματα ταξινόμησης.

## 5. Υβριδικά μοντέλα

### 5.1 Υβριδικό ELM-SVM

Υπάρχουν τρεις διαφορετικές προσεγγίσεις όσον αφορά στην εκπαίδευση των νευρωνικών δικτύων πρόσθιας τροφοδότησης: οι παραδοσιακοί αλγόριθμοι που βασίζονται στην κλίση (*gradient - based algorithms*), οι αλγόριθμοι που βασίζονται στην ελαχιστοποίηση των τετραγώνων (*least -square based algorithms*), και οι αλγόριθμοι που βασίζονται στην κλασική μέθοδο βελτιστοποίησης (*standard optimization method based algorithms*).

Για καλύτερα αποτελέσματα σε προβλήματα ταξινόμησης, εκτός από τις παραλλαγές του κλασικού αλγορίθμου ELM, μπορούν να χρησιμοποιηθούν υβριδικά μοντέλα τα οποία συνδυάζουν τα πλεονεκτήματα της ELM (δηλαδή ενός αλγορίθμου από τη δεύτερη κατηγορία) και των μηχανών διανυσμάτων υποστήριξης (από την τρίτη κατηγορία).

Οι αλγόριθμοι SVM, χρησιμοποιούνται σε πάρα πολλές εφαρμογές και βασίζονται στη μέθοδο της βελτιστοποίησης για να βρουν τη λύση, ελαχιστοποιώντας τα σφάλματα εκπαίδευσης και μεγιστοποιώντας το χάρισμα μεταξύ δύο κλάσεων. Από την άλλη πλευρά, οι αλγόριθμοι ELM αναπτύχθηκαν κυρίως για νευρωνικά δίκτυα με ένα μόνο κρυφό επίπεδο κόμβων, οι παράμετροι των κρυφών νευρωνίων επιλέγονται τυχαία και είναι ανεξάρτητοι των δεδομένων εκπαίδευσης, και ανάλογα υπολογίζονται τα βάρη εξόδου. Ουσιαστικά εστιάζουν στην συνάρτηση προσέγγισης των εφαρμογών, προσπαθώντας να ελαχιστοποιήσουν όχι μόνο το σφάλμα εκπαίδευσης αλλά και τη νόρμα των βαρών εξόδου. Στα προβλήματα ταξινόμησης η μέθοδος εύρεσης βέλτιστης λύσης που βασίζεται στη μέθοδο των ελαχίστων τετραγώνων δεν θεωρείται η πλέον κατάλληλη.

Ένας εύκολος τρόπος υλοποίησης υβριδικού μοντέλου (ELM-SVM) για καλύτερες αποδόσεις σε προβλήματα ταξινόμησης είναι η αντικατάσταση των SVM πυρήνων με ELM πυρήνες. Επιπλέον, ο ELM μπορεί να επεκταθεί γραμμικά σε SVM και μάλιστα με λιγότερους περιορισμούς και έχοντας σαν αποτέλεσμα ένα πιο απλοϊκό μοντέλο. Ο

ένας περιορισμός που παύει να υφίσταται είναι η ύπαρξη πολώσεων  $b$  που συναντάμε στο SVM μοντέλο. Η κεντρική ιδέα του αλγορίθμου SVM είναι πως τα δεδομένα εκπαίδευσης αντιστοιχίζονται σε κάποιο χώρο και υπάρχουν πολυεπίπεδα που μπορούν να χωρίζουν σε κατηγορίες τα δεδομένα αυτά με το μέγιστο περιθώριο. Είναι λογικό αυτά τα πολυεπίπεδα να μην διέρχονται απαραίτητα από τον αρχικά ορισμένο χώρο του SVM και για αυτό το λόγο απαιτείται μία πόλωση  $b$  για τη βελτιστοποίηση του περιορισμού

$$t_i(\mathbf{w}^* \boldsymbol{\varphi}(\mathbf{x}_i) + \mathbf{b}) \geq 1 - \xi_i \quad (26)$$

Από τον παραπάνω περιορισμό και από την συνθήκη για την ελαχιστοποίηση σφαλμάτων, προκύπτει ο λόγος για τον οποίο στα SVM μοντέλα πρέπει να τηρείται ο περιορισμός  $\sum_{i=1..N} \alpha_i t_i = 0$  και  $0 \leq \alpha_i \leq C$ . Αντίθετα με τα μοντέλα SVM, στα μοντέλα ELM τα διαφορετικά πολυεπίπεδα βασικά διέρχονται από τον αυθεντικό χώρο που έχει οριστεί συνεπώς μένει να ικανοποιηθεί μόνο ο περιορισμός  $0 \leq \alpha_i \leq C$ . Με άλλα λόγια η λύση που προσφέρει το SVM σύστημα είναι μία υπό-λύση από το σύνολο των λύσεων που δίνει ο ELM. Ο ELM ψάχνει τη βέλτιστη λύση μέσα στον κύβο  $[0, C]^N$ , ενώ ο SVM ψάχνει τη λύση στο πολυεπίπεδο  $\sum_{i=1..N} \alpha_i t_i = 0$  μέσα στον κύβο  $[0, C]^N$ . Η προσπάθεια που γίνεται στο μοντέλο του ELM να ελαχιστοποιηθεί η νόρμα των βαρών εξόδου αντιστοιχίζεται με την ιδιότητα των SVM's να χωρίζουν δύο κλάσεις με το μέγιστο δυνατό περιθώριο.

Συνοψίζοντας, το υβριδικό μοντέλο (SVM-ELM) χρησιμοποιεί, σύμφωνα με τον αλγόριθμο ELM, τυχαία συνάρτηση αντιστοίχισης  $h(x)$ , αφού όλοι οι παράμετροι της επιλέγονται τυχαία σε αντίθεση με ότι υπαγορεύει η λογική των SVM's, και από την άλλη σύμφωνα με το μοντέλο SVM χρησιμοποιείται η κλασική μέθοδος βελτιστοποίησης για την επίλυση του προβλήματος (αντί για τη μέθοδο των ελαχίστων τετραγώνων που χρησιμοποιεί ο ELM). Παύει να ισχύει ο περιορισμός  $\sum_{i=1..N} \alpha_i t_i = 0$  καθώς και η παρουσία των πολώσεων  $b$ . Με αυτόν τον τρόπο προσαρμόζεται γραμμικά ο αλγόριθμος ELM στον SVM, το νέο μοντέλο έχει ως κύριο χαρακτηριστικό τη συνεκτικότητα και επιτυγχάνεται καλύτερη απόδοση γενίκευσης και καλύτεροι χρόνοι ελέγχου και εκπαίδευσης συγκριτικά με τους αλγορίθμους SVM, ενώ τα αποτελέσματα συγκριτικά με τους αλγορίθμους ELM είναι παρόμοια. Το συγκεκριμένο μοντέλο ουσιαστικά εξαλείφει το μειονέκτημα που έχουν οι SVM ταξινομητές όσον αφορά στην υπερπροσαρμογή των δεδομένων.

## 5.2 Υβριδικό ELM- GMC

Ένα κλασικό πρόβλημα ταξινόμησης είναι η ταξινόμηση απλής κλάσης (*Single Class Classification – SCC*). Ουσιαστικά ξεχωρίζεται μία κλάση δεδομένων τα οποία αναφέρονται ως θετικά, από τα υπόλοιπα δεδομένα τα οποία αναφέρονται ως αρνητικά. Τα προβλήματα τέτοιου τύπου είναι ευρέως διαδεδομένα, όπου τα θετικά δεδομένα είναι εύκολα διαθέσιμα, αντιθέτως τα αρνητικά είναι πολύ ακριβό και δύσκολο να αποκτηθούν (Για παράδειγμα τα δεδομένα για τους υγιείς ανθρώπους είναι διαθέσιμα

πολύ εύκολα, αλλά για αυτούς που υποφέρουν από κάποια πάθηση, ύστερα από σειρά εξετάσεων είναι αρκετά ακριβά. Από την άλλη τα δεδομένα των ασθενών βρίσκονται σχεδόν όλα συγκεντρωμένα σε ένα νοσοκομείο).

Οι παραδοσιακοί αλγόριθμοι δεν είναι κατάλληλοι για το παραπάνω πρόβλημα αφού ουσιαστικά δεν μπορούν να διατηρήσουν σε μία ισορροπία το όριο μεταξύ θετικής και αρνητικής κλάσης, και επίσης πολύ συχνά δεν εμφανίζεται καν αρνητική κλάση. Συνεπώς υστερούν πολύ και σε ακρίβεια. Μία άλλη μέθοδος προσέγγισης είναι να

οριστούν τα όρια (κλειστά ή ανοιχτά) μιας υπέρ-σφαίρας η οποία να περικλείει όλα τα θετικά δεδομένα . Το σημείο κλειδί εδώ, είναι να οριστεί όσο το δυνατόν καλύτερα και ακριβέστερα το όριο της σφαίρας που ουσιαστικά θα περικλείει τα θετικά και δεν θα έχει κανένα κοινό σημείο με τα αρνητικά δεδομένα. Ο αλγόριθμος MC (*Mapping Convergence*) προσπαθεί ουσιαστικά να βρει αυτό το όριο εκμεταλλευόμενος το φυσικό χώρισμα μεταξύ της θετικής και της αρνητικής κλάσης.

Ο αλγόριθμος MC χωρίζεται σε δύο φάσεις. Αρχικά χρησιμοποιεί το σύνολο των θετικών δεδομένων  $P$  με το σύνολο  $U$  για να ορίσει έναν προσχέδιο αλγόριθμο ταξινόμησης  $\psi_1$ . Ο συγκεκριμένος αλγόριθμος, στη συνέχεια, εφαρμόζεται στο σύνολο  $U$ , για να μπορέσει να δημιουργήσει το «ισχυρό» αρνητικό σύνολο  $N_0$ . Με τον όρο ισχυρό εννοούνται όλα τα αρνητικά δεδομένα τα οποία απέχουν πολύ από την περιοχή των θετικών δεδομένων που ορίζεται από το σύνολο  $P$ . Τα δεδομένα τα οποία δεν χαρακτηρίστηκαν ως αρνητικά από τον  $\psi_1$  (και συνεπώς υπονοείται ότι είναι θετικά παραδείγματα) και αποτελούν το σύνολο  $U'$  χρησιμοποιούνται στην επόμενη φάση. Αξίζει να σημειωθεί ότι  $U = U' + N_0$ . Στην συνέχεια, εφαρμόζεται ένας δεύτερος αλγόριθμος  $\psi_2$  στο σύνολο  $P$  και στο σύνολο των αρνητικών δεδομένων  $N$  (με αρχική τιμή  $N_0$ ) ώστε να κατασκευαστεί ένας νέος ταξινομητής ο οποίος με τη σειρά του θα εφαρμοστεί στο σύνολο  $U'$ .

Οι αλγόριθμοι SVMC και GMC είναι παραλλαγές του βασικού MC αλγόριθμου. Στον πρώτο ύστερα από το τέλος κάθε επανάληψης, το σύνολο των δεδομένων εκπαίδευσης επαναπροσδιορίζεται. Στον δεύτερο, υπάρχει ένα κριτήριο παύσης του αλγόριθμου που βασίζεται στην απότομη μείωση του αριθμού των αρνητικών δειγμάτων στο σύνολο  $N_t$ . Όλοι οι αλγόριθμοί που προαναφέρθηκαν χρησιμοποιούν σαν αλγόριθμο  $\psi_2$  τον SVM, γεγονός που τους κάνει αρκετά αργούς στην εκπαίδευση και στον υπολογισμό.

Ένα υβριδικό μοντέλο, το οποίο έχει πολύ καλό ρυθμό ταξινόμησης δημιουργείται όταν στο γενικό πλαίσιο του GMC, χρησιμοποιείται σαν αλγόριθμος  $\psi_2$  ο αλγόριθμος ELM. Με την υλοποίηση του μοντέλου αυτού αποδεικνύεται ότι ο αλγόριθμος  $\psi_2$  δεν είναι απαραίτητο να είναι αλγόριθμος που μεγιστοποιεί το περιθώριο μεταξύ των δύο κλάσεων (*margin maximize algorithm*). Επιπρόσθετα στο συγκεκριμένο μοντέλο το πρώτο κριτήριο παύσης ικανοποιείται όταν οριστεί το «κενό» μεταξύ θετικής και αρνητικής κλάσης ( $N_t$  δηλαδή να είναι κενό σύνολο). Το δεύτερο κριτήριο που χρησιμοποιείται είναι ίδιο με αυτό που προαναφέρθηκε για τον αλγόριθμο GMC και είναι ουσιαστικά και αυτό το οποίο τον διαφοροποιεί από τον αλγόριθμο MC.

Συγκρινόμενο με άλλες μεθόδους ταξινόμησης (IELM, ISVM, OSVM, GMC, ELM\_NN, SVM\_NN), το συγκεκριμένο υβριδικό μοντέλο έχει λίγο υψηλότερο ρυθμό ταξινόμησης συγκριτικά με IELM και ISVM, ενώ συγκρινόμενο με τον αλγόριθμο GMC

παρατηρήθηκε ότι έχει τον ίδιο περίπου ρυθμό ταξινόμησης, το υβριδικό μοντέλο όμως έχει πιο καλούς χρόνους εκπαίδευσης.



## 6. Εφαρμογές του αλγόριθμου ELM σε προβλήματα ταξινόμησης.

### 6.1 Μόρια συμπλέγματος ιστοσυμβατότητας

Τα μόρια συμπλέγματος ιστοσυμβατότητας (*Histocompatibility Complex Molecules*) παίζουν πολύ σημαντικό ρόλο στο αναπνευστικό σύστημα. Στους ανθρώπους, τα μόρια αυτά είναι γνωστά ως αντιγόνα ανθρώπινων λευκοκυττάρων (*Human Leukocyte Antigens*). Τα τελευταία αποτελούνται από πεπτιδία και εμφανίζονται στην επιφάνεια των κυττάρων, ώστε να αναγνωρισθούν από τα Τ-κύτταρα του αναπνευστικού συστήματος. Τα HLA γονίδια παρουσιάζουν μεγάλη πολυμορφικότητα, υπάρχουν πάνω από 2000 γνωστοί τύποι. Λόγω αυτής της μεγάλης ποικιλομορφίας είναι αδύνατη η διεξαγωγή πειραμάτων για κάθε ζεύγος HLA-πεπτιδίου. Συνεπώς τα υπολογιστικά μοντέλα που αναπτύχθηκαν, έχουν σαν στόχο την πρόβλεψη για το αν ένα συγκεκριμένο πεπτιδίο μπορεί να συνδέεται με ένα συγκεκριμένο HLA.

Πολλές κατασκευαστικές μέθοδοι με πολύπλοκους υπολογισμούς, έχουν εφαρμοστεί σε HLA που έχουν γνωστές κρυσταλλικές δομές. Οι ακολουθιακές στατιστικές μέθοδοι ταξινομούνται σε δύο μεγάλες κατηγορίες. Στις AIB μεθόδους, όπως τα δεσμευτικά μοτίβα και τα ποσοτικά πρότυπα (τα οποία υποθέτουν ανεξάρτητες κατανομές των υπολειμμάτων των πεπτιδίων μέσα στο ζεύγος πεπτιδίου-HLA) και στις γενικές μεθόδους όπως η ταξινόμηση, τα δένδρα παλινδρόμησης, τα νευρωνικά δίκτυα, τα μοντέλα Markov και τα δένδρα αποφάσεων, οι οποίες λαμβάνουν υπόψη τους τις επιρροές των διαφορετικών υπολειμμάτων πεπτιδίων σε διαφορετικά σημεία του πεπτιδίου. Τα νευρωνικά δίκτυα, και πιο συγκεκριμένα οι μέθοδοι ταξινόμησης δίνουν πολύ καλά και γρήγορα αποτελέσματα όσον αφορά τις προβλέψεις. Σαν αλγόριθμος ταξινόμησης με όλα τα χαρακτηριστικά που προαναφέρθηκαν έχει χρησιμοποιηθεί και ο ELM με τρία επίπεδα (είσοδος – έξοδος και κρυφό), οι παράμετροι επιλέγονται τυχαία και η έξοδος υπολογίζεται βάση της αντιστροφής της μήτρας H. Οι χρόνοι εκπαίδευσης είναι πολύ καλύτεροι σε σχέση με τον κλασικό αλγόριθμο πίσω διάδοσης του λάθους (BP) και όσο αυξάνεται ο αριθμός των κρυφών κόμβων τόσο μεγαλύτερη ακρίβεια επιτυγχάνεται και επίσης αντιμετωπίζεται αποτελεσματικά το φαινόμενο της υπερπροσαρμογής. Η πρόβλεψη αυτή λαμβάνεται από έναν συνδυασμό συνεισφοράς μεμονωμένων αμινοξέων σε συγκεκριμένα σημεία του πεπτιδίου και συνεισφορές της αλληλεπίδρασης της πλευρικής αλυσίδας με την κεντρική.

### 6.2 Ταξινόμηση ψυχικών καταστάσεων από ηλεκτροσήματα του εγκεφάλου

Η ανάπτυξη του Brain Computer Interface (BCI) έχει σαν στόχο την παροχή επικοινωνίας και ελέγχου δυνατοτήτων σε άτομα με δυσλειτουργίες. Πρόκειται για μία διεπιστημονική έρευνα με τη συμμετοχή εμπειρογνομώνων νευρολογίας, ψυχολογίας, μαθηματικών, μηχανικών και προγραμματιστών. Οποιαδήποτε πρακτική εφαρμογή της αρχιτεκτονικής BCI απαιτεί την προ-επεξεργασία σήματος, τον προσδιορισμό χαρακτηριστικών και την ταξινόμηση. Ένας ιδανικός ταξινομητής για το BCI πρέπει να

είναι σε θέση να προσφέρει υψηλή ακρίβεια στην ταξινόμηση με την ελάχιστη δυνατή πολυπλοκότητα στους υπολογισμούς.

Οι ταξινομητές για την επίλυση των συγκεκριμένων προβλημάτων μπορούν να χωριστούν σε γραμμικούς και μη γραμμικούς. Επιπρόσθετα κάθε ένας από αυτούς εμπεριέχει στατιστικές τεχνικές, όπως είναι η γραμμική διακρίνουσα ανάλυση (*Linear Discriminant Analysis LDA*), ο κρυφός ταξινομητής Markov (*Hidden Markov Classifier*) και η z κλίμακας διακρίνουσα ανάλυση (*z-scale base Discriminant Analysis ZDA*). Το βασικό μειονέκτημα των τεχνικών αυτών είναι ότι δε δουλεύουν καλά για μη γραμμικά προβλήματα. Οι μηχανές διανυσμάτων υποστήριξης και τα δίκτυα οπίσθιας τροφοδότησης με τη σειρά τους βελτίωσαν την ακρίβεια της ταξινόμησης αλλά οι χρόνοι τους είναι πολύ αργοί και η πολυπλοκότητα τους μεγάλη. Ο αλγόριθμος ELM με την υψηλή απόδοση του δίνει ουσιαστικά λύση σε πολλά προβλήματα που αντιμετωπίζουν οι ταξινομητές που προαναφέρθηκαν (έχει πολύ καλύτερους χρόνους εκπαίδευσης και ελέγχου).

Τα μετρήσιμα ηλεκτρικά σήματα του εγκεφάλου, όπως το ηλεκτροεγκεφαλογράφημα, μεταφέρουν πληροφορίες σχετικά με διανοητικές δραστηριότητες. Το πρόβλημα προς λύση είναι η αναγνώριση των διανοητικών εργασιών και συνεπώς η ταξινόμηση των ηλεκτροεγκεφαλικών σημάτων στις διαφορετικές κλάσεις, όπου η καθεμιά αντιπροσωπεύει μία διανοητική διεργασία. Μία πολύ γνωστή βάση δεδομένων ηλεκτροεγκεφαλικών σημάτων που έχει χρησιμοποιηθεί για πολλά πειράματα και ελέγχους, αποτελείται από σήματα που προκύπτουν ύστερα από λεπτομερή καταγραφή και αντιπροσωπεύουν πέντε βασικές διανοητικές δραστηριότητες: χαλάρωση, σύνταξη επιστολής, πολλαπλασιασμός, μέτρηση και περιστροφή ανά συνεδρία.

Τα ακατέργαστα δεδομένα, όπως προαναφέρθηκε, αποτελούνται από πέντε βασικές διανοητικές καταστάσεις. Αρχικά τα σήματα ομαλοποιούνται με μέση τιμή μηδέν και διασπορά ένα. Στη συνέχεια γίνεται εύρεση και εξαγωγή χαρακτηριστικών από αυτά τα

ακατέργαστα δεδομένα. Τα ηλεκτροεγκεφαλικά σήματα υπόκεινται αλλαγές όσον αφορά στο πλάτος τους και τη συχνότητά τους ανάλογα με την διανοητική δραστηριότητα. Τα διανύσματα αυτών των χαρακτηριστικών χρησιμοποιούνται για να εκπαιδεύσουν τους ταξινομητές. Τέλος οι ταξινομητές αυτοί ελέγχονται με δεδομένα τα οποία δεν είχαν εμφανιστεί κατά τη διάρκεια της εκπαίδευσης ώστε να διαπιστωθεί η ακρίβεια τους.

Ο αλγόριθμος ELM χρησιμοποιείται για την ταξινόμηση στο παραπάνω πρόβλημα και συγκρίνονται τα αποτελέσματα του με τον αλγόριθμο SVM και τον αλγόριθμο BPNN. Παρατηρήθηκε ότι η ακρίβεια ταξινόμησης είναι παρόμοια και στους τρεις αλγόριθμους. Επιπλέον ο χρόνος αναζήτησης παραμέτρων του βέλτιστου ταξινομητή είναι αισθητά χαμηλότερος στον αλγόριθμο ELM.

### 6.3 Ταξινόμηση με ηλεκτρονική μύτη

Παρόμοια εφαρμογή με την παραπάνω συναντάται και στην περίπτωση της ταξινόμησης δεδομένων ηλεκτρονικής μύτης. Οι καφέδες του εμπορίου είναι ουσιαστικά μείγμα από καφέ διαφόρων προελεύσεων των οποίων η ανάλυση και ο έλεγχος τους έχει μεγάλη σημασία. Ο χαρακτηρισμός του κάθε καφέ γίνεται χρησιμοποιώντας τη χημική ανάλυση ενός κλάσματος του, όπως το επάνω τμήμα των πράσινων ή των ψητών κόκκων ή τα φαινολικά κλάσματα. Ως τώρα έχουν αναγνωρισθεί τουλάχιστον 700 διαφορετικά είδη καφέ βάση του τύπου, της προέλευσης και της παρασκευής τους. Κανένα από τα μόρια αυτά δεν μπορεί να ταυτοποιηθεί ως «δείκτης», από την άλλη στις περισσότερες περιπτώσεις πρέπει να εξετασθεί ολόκληρο το φάσμα, όπως για παράδειγμα με την αέρια χρωματογραφία.

Η ηλεκτρονική μύτη, βασιζόμενη στην αέρια χρωματογραφία, και συνδυάζοντας το κατάλληλο λογισμικό βοηθά να προσδιοριστεί η συγκέντρωση των αέριων συστατικών στο μείγμα και με αυτό τον τρόπο μπορεί να βρίσκει τις διαφορές και άρα διαχωρίζει τα μείγματα αυτά. Ο αλγόριθμος ταξινόμησης που έχει χρησιμοποιηθεί για αυτήν την κατηγοριοποίηση είναι ο RI-ELM ο οποίος επιτυγχάνει ακρίβεια της τάξης του 90%.

### 6.4 Ταξινόμηση πρωτεϊνών

Πολλά είναι τα βιολογικά δεδομένα τα οποία συλλέγονται σε συγκεκριμένες βάσεις δεδομένων και ασχολούνται με την αλληλουχία των γονιδίων στην αλυσίδα του γενετικού υλικού. Οι πρωτεϊνικές αλληλουχίες, γνωστές και ως πολυπεπτίδια είναι οργανικές ενώσεις που κατασκευάζονται από αμινοξέα διευθετημένα σε μία γραμμική αλυσίδα και διπλωμένα σε μία σφαιρική μορφή. Τα αμινοξέα σε μια αλυσίδα πολυμερούς ενώνονται μεταξύ τους με πεπτιδικούς δεσμούς μεταξύ των καρβοξυλίου και αμινο ομάδων γειτονικών υπολειμμάτων αμινοξέων. Η αλληλουχία των αμινοξέων σε μια πρωτεΐνη ορίζεται από την αλληλουχία ενός γονιδίου, η οποία κωδικοποιείται στο γενετικό κώδικα. Ένα γονίδιο είναι οποιοδήποτε τμήμα κατά μήκος του δεοξυριβονουκλεϊκού οξέος (DNA) που κωδικοποιεί οδηγίες οι οποίες επιτρέπουν σε ένα κύτταρο την παραγωγή ενός συγκεκριμένου προϊόντος. Τυπικά, μία πρωτεΐνη όπως ένα ένζυμο ενεργοποιεί μία συγκεκριμένη δράση.

Η ταξινόμηση των πρωτεϊνών έχει σαν απώτερο στόχο την πρόβλεψη της δομής και της λειτουργίας των νέων πρωτεϊνών. Μία πρωτεΐνη μπορεί να ταξινομηθεί σε συγκεκριμένη οικογένεια πρωτεϊνών οι οποίες έχουν κάποια κοινά χαρακτηριστικά. Γενικά, δύο αλληλουχίες πρωτεϊνών κατατάσσονται στην ίδια κατηγορία, εάν τα μοτίβα των χαρακτηριστικών τους που εξάγονται από ακολουθιακούς αλγόριθμους δείχνουν υψηλή ομολογία. Το κέρδος από αυτήν την ταξινόμηση φαίνεται στην ανάλυση των διάφορων μορίων, η οποία μπορεί να γίνει μέσα σε ένα ευρύτερο πλαίσιο (την οικογένεια πρωτεϊνών) και όχι για καθεμία μεμονωμένη πρωτεΐνη χωριστά.

Υπάρχουν πολλές προσεγγίσεις για αυτό το πρόβλημα ταξινόμησης, όπως στατιστικές τεχνικές, δένδρα αποφάσεων και νευρωνικά δίκτυα. Τα νευρωνικά δίκτυα υπερέχουν έναντι των υπόλοιπων μεθόδων για δύο κύριους λόγους: οι ιδιότητες των πρωτεϊνικών

αλληλουχιών κατανέμονται σε ένα πολυδιάστατο χώρο με πολύπλοκα χαρακτηριστικά, πράγμα που είναι δύσκολο να μοντελοποιηθεί βάση των μεθόδων που στηρίζονται σε παραμέτρους. Επιπλέον οι μέθοδοι που βασίζονται στα δένδρα αποφάσεων δεν παρουσιάζουν κανένα πλεονέκτημα στην ταξινόμηση μοτίβων με συνεχή χαρακτηριστικά, ειδικά όταν ο αριθμός των ιδιοτήτων είναι μεγαλύτερος. Τα νευρωνικά δίκτυα είναι σε θέση να επεξεργάζονται τυχαία δεδομένα τα οποία τροφοδοτούν το δίκτυο με συνεχή ρυθμό.

Ο κλασικός αλγόριθμος ELM έχει χρησιμοποιηθεί για την ταξινόμηση τέτοιων προβλημάτων και τα αποτελέσματά του έχουν συγκριθεί με τους κλασικούς ταξινομητές. Προκύπτει ότι η ακρίβεια της ταξινόμησης που κάνει, είναι ελάχιστα καλύτερη από τον αλγόριθμο BP είναι αρκετά πιο εύκολο όμως να εφαρμοστεί. Επιπρόσθετα, πολλά πειράματα έχουν γίνει και με αρκετές παραλλαγές του αλγορίθμου ELM ( VOP-ELM, OP-ELM, V-ELM) οι οποίοι έχουν συγκριθεί με τον SVM και τον BP. Οι αλγόριθμοι VOP-ELM και V-ELM, αναπτύχθηκαν συγκεκριμένα για την κατανομή των πρωτεϊνών σε οικογένειες. Ο αλγόριθμος VOP-ELM επιτυγχάνει την αποδοτικότερη ταξινόμηση και τους καλύτερους χρόνους εκπαίδευσης συγκριτικά με όλους τους υπόλοιπους αλγορίθμους, ενώ ο V-ELM ξεπερνά σε ακρίβεια ταξινόμησης τον κλασικό ELM και BP, αλλά οι χρόνοι εκπαίδευσης του είναι παρόμοιοι με τους παραπάνω.

## 6.5 Ταξινόμηση ειδών μουσικής

Η ψηφιακή μουσική στις μέρες μας είναι ευρέως διαδεδομένη και ο όγκος των δεδομένων πολύ μεγάλος σε αντίθεση με τα μέσα κατηγοριοποίησης της ψηφιακής μουσικής όπου θα χαρακτηρίζονταν φτωχά. Τα δεδομένα που βρίσκονται στην ίδια

κατηγορία μοιράζονται κάποια κοινά χαρακτηριστικά, ενώ μπορεί από κάποια άλλη οπτική να θεωρούνται ανεξάρτητα. Η ταξινόμηση της ψηφιακής μουσικής σε διάφορα είδη βελτιώνει αυτόματα την αναζήτηση και την αποθήκευση των δεδομένων.

Ένα άλλο πρόβλημα είναι ότι η εκμετάλλευση της ψηφιακής μουσικής είναι υπολογιστικά ακριβή. Οι μέθοδοι με τις οποίες εξάγονται πληροφορίες, όπως η ανάλυση ηχοχρώματος, είναι κατάλληλη μόνο στο να δώσει μία εκτίμηση συνεπώς απουσιάζει η ακρίβεια που χρειάζεται για την ταξινόμηση σε διάφορα μουσικά είδη. Τα χαρακτηριστικά των χαμηλής έντασης σημάτων, όπως η ενέργεια, μπορούν εύκολα να υπολογιστούν με κάποιους αλγόριθμους αλλά δε μεταφράζεται σε ανθρώπινη αντίληψη εύκολα και συνεπώς αυτό επηρεάζει την κατηγοριοποίηση.

Στις διάφορες μεθόδους ταξινόμησης έχουν χρησιμοποιηθεί χαρακτηριστικά που σχετίζονται με τη μουσική επιφάνεια, τη φασματική ροή, τις μηδενικές διασταυρώσεις, ενώ τα χαρακτηριστικά του ρυθμού και της χαμηλής ενέργειας προέρχονται από ένα ιστόγραμμα παλμού το οποίο συμπεριλαμβάνει την περιοδικότητα και τα σχετικά πλάτη των κορυφών καθώς και τις αναλογίες αυτών. Για παράδειγμα ο Gauss -ιανός

ταξινομητής επιτυγχάνει ακρίβεια 62% με έξι διαφορετικά είδη μουσικής. Η ταξινόμηση έγινε με βάση το ύψος, τη διάθεση, το συναίσθημα, την πολυπλοκότητα και το τέμπο. Άλλες μέθοδοι για παράδειγμα η KNN επιτυγχάνει 96% ακρίβεια.

Τα χαρακτηριστικά που προαναφέρθηκαν είναι ζωτικής σημασίας για την απόδοση της ταξινόμησης καθώς θα πρέπει να αντιπροσωπεύουν το τραγούδι, όσο πιο χαρακτηριστικά γίνεται. Τα χαρακτηριστικά αυτά πρέπει να προσδιοριστούν και στο πεδίο του χρόνου και στην συχνότητα. Αν θεωρηθεί ότι το μουσικό σήμα συμβολίζεται με  $x$  και  $x_r$  το  $r$ -ιοστό μη επικαλυπτόμενο μουσικό πλαίσιο που έχει κατασκευαστεί από  $N$  επιλεγμένα δείγματα ισχύει

$$x_r(n) = \begin{cases} x(n+r_N), & 0 \leq n \leq (N-1) \quad (27) \\ 0 & \text{για οποιαδήποτε άλλη τιμή του } n \end{cases}$$

Παρακάτω ορίζονται συνοπτικά τα βασικά χαρακτηριστικά τα οποία έχουν προαναφερθεί

Ρυθμός μηδενικής διασταύρωσης (*zero-crossing rate*): Είναι οι φορές που το σήμα αλλάζει πρόσημο μέσα σε ένα συγκεκριμένο χρονικό διάστημα. Μπορεί να χρησιμοποιηθεί ως συχνότητα του σήματος, όσο πιο υψηλό είναι το νούμερο τόσο μεγαλύτερη η συχνότητα του σήματος και υπολογίζεται από τον παρακάτω τύπο.

$$z_r = \sum |s[x_r(n)] - s[x_r(n-1)]| \quad (28) \quad \text{και}$$

$$s[x_r(n)] = \begin{cases} 1, & x_r(n) \geq 0 \quad (29) \\ 0, & x_r(n) < 0 \end{cases}$$

Βραχυπρόθεσμη ενέργεια (*short-term energy*): Η ενέργεια συνήθως χρησιμοποιείται σαν μία εκτίμηση της έντασης του ήχου ενός μουσικού πλαισίου. Ένα σήμα χαμηλής ενέργειας χαρακτηρίζει «αραιά» μουσικά είδη, όπως instrumental πιάνο και κλασική μουσική αφού έχουν μεγαλύτερα χρονικά διαστήματα σιωπής. Η ενέργεια δίδεται από τον τύπο

$$E_r = \sum |x_r(n)|^2 \quad (30)$$

Μέση τετραγωνική ρίζα (*root mean square*) : Είναι το μέτρο της «δύναμης» του σήματος

$$x_r - \text{rms} = ((1/N) * \sum |x_r(n)|^2)^{1/2} \quad (31)$$

Συντελεστής κορυφής (*crest factor*) : ορίζεται ως ο λόγος κορυφής-RMS ενός πλαισίου. Είναι ένα χαρακτηριστικό που υπολογίζει την οξύτητα του πλαισίου. Ο συντελεστής κορυφής, μπορεί να υπολογιστεί από:

$$CF_r = x_r\text{-peak} / x_r\text{-rms} \quad (32)$$

Φασματικός κεντροειδής (*spectral centroid*): είναι η διάμεση συχνότητα ενός σήματος. Μπορεί να θεωρηθεί μέτρο της φασματικής φωτεινότητας του συγκεκριμένου σήματος.

Ορίζεται ως

$$SCr = (\sum k \times X_r(k)) / (\sum X_r(k)) \quad (33)$$

Φασματική ροή (*Spectral Flux*): είναι ένα μέτρο της φασματικής μεταβολής μεταξύ δύο διαδοχικών καρέ μουσικής

$$SF = (\sum (|X_r(k)| - |X_r(k-1)|)^2)^{1/2} \quad (34)$$

*A - Frequency Cepstral Coefficients*: είναι μέθοδος ανάλυσης του FFT του πλαισίου μουσικής λόγω της ομοιότητας του με το ανθρώπινο μοντέλο ακοής. Περιλαμβάνει έξι στάδια για εξαγωγή χαρακτηριστικών.

Ειδική αίσθηση εντάσεως (*Specific Loudness Sensation*): είναι ένα μέτρο της ανθρώπινης αντίληψης για την ένταση του ήχου ανά ζώνη συχνοτήτων (κλίμακα Bark)

Για το παραπάνω πρόβλημα ταξινόμησης του όγκου μουσικών δεδομένων σε διαφορετικά είδη μπορεί να χρησιμοποιηθεί ο αλγόριθμος ELM. Συγκρινόμενος δε με τον αλγόριθμο SVM για το ίδιο πρόβλημα προκύπτει ότι έχουν παρόμοια ακρίβεια ταξινόμησης χωρίς κάποιος να υπερτερεί ξεκάθαρα. Τα ποσοστά ακρίβειας του ELM αγγίζουν το 85%, ενώ του SVM το 82%.

## 6.6 Αναγνώριση προσώπων – Ταξινόμηση αντικειμένων –Αναγνώριση δραστηριοτήτων

### 6.6.1 Πρόσωπα

Η αναγνώριση προσώπων έχει απασχολήσει αρκετά την ερευνητική κοινότητα τις τελευταίες δεκαετίες λόγω του ότι παρουσιάζεται σε αρκετούς τομείς η ανάγκη για αναγνώριση φυσιογνωμιών (δίπλωμα οδήγησης, διαβατήρια, διαδραστικότητα μεταξύ

ανθρώπου και υπολογιστή κτλ). Η αναγνώριση προσώπου είναι μη παρεμβατική, δηλαδή η φωτογραφία βγαίνει και από εκεί και πέρα μπορεί να ταυτοποιηθεί ή να επαληθευτεί χωρίς κάποια ιδιαίτερη γνώση ή φυσική αλληλεπίδραση με το αντικείμενο. Επιπλέον δεν χρειάζεται κάποιος εμπειρογνώμονας για να ερμηνεύσει και να αναλύσει τα δεδομένα, τα οποία συλλέγονται εύκολα με απλή χρήση συσκευών απεικόνισης.

Οι άνθρωποι αναγνωρίζουν άλλα πρόσωπα με μεγάλη ευκολία, ωστόσο μία αυτοματοποιημένη αναγνώριση προτύπου είναι αρκετά πολύπλοκη λόγω του ότι τα πρόσωπα ανήκουν σε μία κατηγορία φυσικών αντικειμένων όπου τα διάφορα χαρακτηριστικά δεν μπορούν να αντιστοιχιστούν με γεωμετρικά χαρακτηριστικά. Το πλεονέκτημα του ηλεκτρονικού αυτού συστήματος, από την άλλη, έχει το πλεονέκτημα να μπορεί να αναγνωρίσει απεριόριστα πρόσωπα ενώ ο ανθρώπινος νους έχει περιορισμένη μνήμη.

Το αυτοματοποιημένο σύστημα αναγνώρισης προσώπων περιλαμβάνει τρία στάδια: ανίχνευση προσώπου, εξαγωγή χαρακτηριστικών και αναγνώριση προσώπου. Πολλοί ερευνητές προς αυτήν την κατεύθυνση εισήγαγαν την έννοια του *localized face* για την εξαγωγή χαρακτηριστικών. Αυτό από την μία δίνει μεγαλύτερη ακρίβεια, από την άλλη αυξάνει την πολυπλοκότητα των διάφορων υπολογισμών. Μόλις εντοπιστούν τα σημεία ενδιαφέροντος η διαδικασία απλοποιείται κατά ένα μεγάλο μέρος εφόσον κάποιες ασαφείς και κάποιες λάθος πληροφορίες αποβάλλονται.

Η αναγνώριση προσώπων χωρίζεται σε δύο μεγάλες κατηγορίες. Η κυριότερη σχετίζεται άμεσα με τα κύρια στοιχεία ενός ανθρώπινου προσώπου όπως η μύτη, τα μάτια, το στόμα κτλ. Είναι αρκετά περίπλοκη η εξαγωγή χαρακτηριστικών γιατί οι ομοιότητες των στοιχείων αυτών είναι πολύ μεγάλες και οι διαφορές ανεπαίσθητες. Για το λόγο αυτό αναπτύχθηκαν αρκετές μέθοδοι που έχουν σαν σκοπό τη μείωση του αριθμού διαστάσεων του χώρου πάνω στον οποίο εφαρμόζεται ο αλγόριθμος ταξινόμησης. Αρχικά μία από τις πιο διαδεδομένες τεχνικές αναπαριστά τα ανθρώπινα πρόσωπα σαν ένα γραμμικό συνδυασμό σταθμισμένων ιδιοδιανυσμάτων χρησιμοποιώντας την ανάλυση σε κύριες συνιστώσες (*principal component analysis PCA*). Επιπλέον αναπτύχθηκαν η ανάλυση ανεξάρτητων συνιστωσών (*independent Component analysis ICA*), η γραμμική διακρίνουσα ανάλυση (*Linear Discriminant Analysis LDA*) και άλλοι. Οι μέθοδοι που βασίζονται στους μετασχηματισμούς βελτιώνουν πολύ την απόδοση των αυτοματοποιημένων συστημάτων αναγνώρισης.

Μία μέθοδος για αναγνώριση η οποία έχει προταθεί, είναι ο συνδυασμός του κλασικού αλγορίθμου ELM και της μεθόδου PCA. Οι εικόνες από τη βάση δεδομένων μετατρέπονται σε εικόνες κλίμακας του γκρι, πράγμα που έχει σαν αποτέλεσμα την μείωση του μεγέθους τους στο υποδιπλάσιο. Κάθε βάση δεδομένων χωρίζεται στα σύνολα εκπαίδευσης και τα σύνολα ελέγχου με τέτοιο τρόπο ώστε το 40-45% των φωτογραφιών ενός υποκειμένου να βρίσκεται στο πρώτο σύνολο και το υπόλοιπο στο δεύτερο. Ο αρχικός μετασχηματισμός (*curvelet*) χρησιμοποιείται για να δημιουργήσει αρχικά διανύσματα χαρακτηριστικών, δεδομένου ότι προσφέρει υψηλότερη απόδοση με την παρουσία ιδιομορφιών σε χώρο υψηλών διαστάσεων λόγω του ότι ενισχύει τα

στοιχεία υψηλότερων συχνοτήτων στα οποία οι λάθος επιρροές ελαχιστοποιούνται. Η αποσύνθεση των φωτογραφιών γίνεται βάση του μετασχηματισμού *cirvelet* και υπολογίζεται σε τρεις κλίμακες και οκτώ γωνιακές κατευθύνσεις, δημιουργώντας 25 ξεχωριστές υποζώνες. Στη συνέχεια ο αλγόριθμος B2DPCA, (όπου η ουσιαστική του διαφορά με την μέθοδο PCA είναι ότι βασίζεται σε δισδιάστατες μήτρες από ότι μονοδιάστατες) δημιουργεί μοναδικά σύνολα χαρακτηριστικών και ελαχιστοποιεί την υπολογιστική πολυπλοκότητα του δικτύου. Στη συνέχεια για κάθε σύνολο δεδομένων επιλέγονται αυτά των οποίων έχει μειωθεί ο αριθμός των διαστάσεων ώστε να αποτελέσουν το σύνολο εκπαίδευσης για τον αλγόριθμο ELM. Τα εναπομείναντα στοιχεία του κάθε συνόλου χρησιμοποιούνται σε θέματα διαχωρισμού μέσα στο δίκτυο. Βασική προϋπόθεση αποτελεί το ότι δεν υπάρχει καμία γνώση εξαρχής για το δίκτυο, το φόντο, τα πρόσωπα ή τον φωτισμό. Η προτεινόμενη μέθοδος επιτυγχάνει βελτιωμένη αναγνώριση σε σημαντικά ταχύτερο ρυθμό συγκριτικά με τις υπόλοιπες κλασικές τεχνικές. Επιπλέον είναι ανεξάρτητη από τον αριθμό των πρωτοτύπων που χρησιμοποιούνται για την εκπαίδευση, των επιπέδων της αποσύνθεσης των φωτογραφιών με βάση τον μετασχηματισμό που προαναφέρθηκε και τον αριθμό των κρυφών νευρώνων.

### 6.6.2 Αντικείμενα

Η ίδια ακριβώς διαδικασία μπορεί να πραγματοποιηθεί για την ταξινόμηση αντικειμένων και όχι προσώπων. Η αυτοματοποιημένη διαδικασία ταξινόμησης αντικειμένων είναι επίσης πολύπλοκη καθώς υπάρχουν πολλές εναλλακτικές θέσεις και καταστάσεις του ίδιου αντικειμένου, διαφορετικές συνθήκες φωτογράφισης, φωτεινότητα, οξύτητα κτλ. Και εδώ υπάρχει η εξαγωγή χαρακτηριστικών με διάφορες μεθόδους (χαρακτηριστικά μπορεί να θεωρηθούν το χρώμα, η υφή, το κέντρο βάρους και άλλα). Και πάλι η κάθε εικόνα μετατρέπεται σε εικόνα κλίμακας του γκρι, η μέθοδος που χρησιμοποιείται για την εξαγωγή χαρακτηριστικών είναι η PCA ή η 2D-PCA, και στη συνέχεια εφαρμόζεται ο αλγόριθμος ELM για εκπαίδευση και έλεγχο.

### 6.6.3 Δραστηριότητες

Για την αναγνώριση ανθρώπινων δραστηριοτήτων μπορεί να ακολουθηθεί παρόμοια διαδικασία αλλά αντί για φωτογραφίες η επεξεργασία γίνεται πάνω σε βίντεο. Αρχικά υπολογίζονται τα υβριδικά διανύσματα των χαρακτηριστικών και γίνεται η αλλαγή μεγέθους των πλαισίων του βίντεο και η μετατροπή χρώματος του στην κλίμακα του γκρι. Έπειτα χρησιμοποιείται η μέθοδος 3D DT-CWT για την εξαγωγή των συντελεστών, οι οποίοι περιέχουν πληροφορίες χωροχρονικές σχετικές με τα ογκομετρικά δεδομένα των διαφορετικών κινούμενων αντικειμένων. Για να δημιουργηθούν διαφορετικά και χαμηλών διαστάσεων χωροχρονικά δεδομένα από τα βίντεο, εφαρμόζεται η μέθοδος 2D-PCA στις διάφορες υποζώνες των πολλαπλών αναλύσεων αποσύνθεσης, γεγονός που καταλήγει σε πολύ μικρότερου μεγέθους χαρακτηριστικά διανυσμάτων. Η δεύτερη κατηγορία χαρακτηριστικών, τα τοπικά



χαρακτηριστικά, εξάγονται εφαρμόζοντας τη μέθοδο ASIFT γύρω από σταθερά σημεία ενδιαφέροντος τα οποία ανιχνεύονται χρησιμοποιώντας την Harris Laplace και την Hessian Laplace υιοθετώντας παράλληλα και μία μέθοδο κλαδέματος ώστε να αποκλειστούν τα αποτελέσματα της ASIFT τα οποία δεν έχουν καμία σχέση με την προσδοκώμενη δραστηριότητα. Τέλος κατασκευάζονται οπτικά λεξιλόγια, τα οποία χρησιμοποιούν και τα δύο είδη χαρακτηριστικών, ώστε να αντιστοιχίσουν συγκεκριμένη ταμπέλα η οποία θα χρησιμοποιηθεί όταν το δεδομένο γίνει είσοδος στον αλγόριθμο ELM. Η προσέγγιση μέσω των λεξιλογίων μειώνει αισθητά τους υπολογισμούς που σχετίζονται με μία εικόνα σε όλα τα σύνολα εκπαίδευσης. Αρχικά αναγνωρίζεται ένας μικρός αριθμός ομάδων με αισθητά διαφορετικά χαρακτηριστικά (μεγάλη ανομοιότητα). Στη συνέχεια μπορεί να εφαρμοστεί και ο αλγόριθμος k-means.

### **6.7 Ταξινόμηση μικροσυστοιχιών γονιδίων για τη διάγνωση καρκίνου**

Στο πρόβλημα ταξινόμησης που βασίζεται στην έκφραση του γονιδιακού προφίλ για τη διάγνωση καρκινικών ασθενειών, τα δυαδικά προβλήματα έχουν μελετηθεί πολύ περισσότερο από τα προβλήματα κατηγοριοποίησης σε πολλές κλάσεις. Τις περισσότερες φορές η αντιμετώπιση των τελευταίων είναι πολύ δύσκολη συγκριτικά με τα δυαδικά προβλήματα και η ακρίβεια ταξινόμησης μειώνεται δραματικά όσο αυξάνεται ο αριθμός των κλάσεων. Αντί να αντιμετωπιστεί ένα τέτοιο πρόβλημα απευθείας χρησιμοποιείται συνδυασμός δυαδικών ταξινομητών σε μία βάση One - Versus - All (OVA) ή One -Versus - One (OVO). Αυτό φυσικά έχει ως αποτέλεσμα την αύξηση της πολυπλοκότητας και τον χρόνο εκπαίδευσης.

Ο αλγόριθμος ELM μπορεί να εφαρμοστεί και για το παραπάνω πρόβλημα ταξινόμησης δίδοντας πολύ καλά αποτελέσματα. Τα αποτελέσματα του έχουν συγκριθεί με τον αλγόριθμο SVM-OVO και τον SVM-OVA. Ο αλγόριθμος ELM αντιμετωπίζει το πρόβλημα απευθείας ενώ οι υπόλοιπες μέθοδοι χρησιμοποιούν συνδυασμό δυαδικών ταξινομητών οπότε, όπως προαναφέρθηκε έχουν μεγαλύτερη πολυπλοκότητα και χρόνους εκπαίδευσης. Επιπλέον ο αλγόριθμος ELM επιτυγχάνει μεγαλύτερη ακρίβεια ταξινόμησης αλλά και πιο ισορροπημένη ταξινόμηση.

## 7. Παρουσίαση - ανάλυση προβλήματος ταξινόμησης κτιρίων σε ενεργειακές κλάσεις

### 7.1 Εισαγωγή

Όπως ήδη προαναφέρθηκε, ο αλγόριθμος ELM χρησιμοποιείται κατά κόρον σε προβλήματα ταξινόμησης και βρίσκει εφαρμογή σε πολλά διαφορετικά επιστημονικά πεδία όπως η βιοιατρική, η ρομποτική, η μηχανική ακόμα και η τέχνη (π.χ. ψηφιακή μουσική) και άλλα. Στόχος, είναι η απόδοση της όσο το δυνατόν καλύτερης γενίκευσης και ακρίβειας, σε όσο το δυνατόν καλύτερους χρόνους.

Το πρόβλημα ταξινόμησης που θα παρουσιαστεί και αναλυθεί παρακάτω, ανήκει στο πεδίο της μηχανικής και αφορά καταρχάς στην ταξινόμηση των κτιρίων, βάσει κάποιων χαρακτηριστικών και μετρήσεων από κάποιο κτίριο, στην **ενεργειακή του κλάση** και αφετέρου στην τροφοδότηση του αλγορίθμου ELM με μια σειρά μετρήσεων, που θα οδηγήσει στην «εκπαίδευση» του για τη δυνατότητα πρόβλεψης της κατάταξης και άλλων κτιρίων σε ενεργειακές κλάσεις. Πέρα από τον αλγόριθμο ELM, πραγματοποιούνται πειράματα και με άλλους αλγορίθμους για να μπορέσουν να συγκριθούν τα αποτελέσματα ακρίβειας ταξινόμησης και χρόνων εκπαίδευσης.

Πρόσφατα, υπήρξε ένας σημαντικός όγκος έρευνας σχετικά με το θέμα της ενεργειακής απόδοσης των κτιρίων λόγω της αυξανόμενης ανησυχίας σχετικά με τη σπατάλη ενέργειας και τις δυσμενείς επιπτώσεις της σπατάλης αυτής στο περιβάλλον. Τα κτίρια σε όλες σχεδόν τις ευρωπαϊκές χώρες, έχουν την υποχρέωση να συμμορφώνονται με τις απαιτούμενες προδιαγραφές για την ενεργειακή απόδοση, σύμφωνα με σχετικό ευρωπαϊκό νόμο. Τα τελευταία χρόνια, έπειτα από μελέτες, έχει αποδειχθεί ότι οι ανάγκες ενός κτιρίου σε κατανάλωση ενέργειας είναι ολοένα και αυξανόμενες και η θέρμανση, ο εξαερισμός και ο κλιματισμός είναι το μεγαλύτερο μερίδιο της κατανάλωσης αφού επηρεάζουν άμεσα τις κλιματικές συνθήκες του εσωτερικού χώρου.

Μία λύση για να αντιμετωπιστεί αυτή η διαρκώς αυξανόμενη ζήτηση σε ενέργεια, είναι η κατασκευή ενεργειακά αποδοτικών κτιρίων με βελτιωμένες ιδιότητες εξοικονόμησης ενέργειας.

Στη συνέχεια αναλύονται κάποιοι όροι, οι οποίοι αναφέρονται συχνά στα επόμενα κεφάλαια – υποκεφάλαια και βοηθούν στην ευκολότερη κατανόηση του προβλήματος ταξινόμησης που επιλέχθηκε.

### 7.1.1 Ορισμοί

- Εξοικονόμηση ενέργειας: Οποιαδήποτε προσπάθεια, με την οποία επιτυγχάνεται περιορισμός της σπατάλης των ενεργειακών αποθεμάτων, με ταυτόχρονη κάλυψη των εκάστοτε ενεργειακών αναγκών.
- Ενεργειακή απόδοση κτιρίου: Η ποσότητα ενέργειας που πράγματι καταναλώνεται ή εκτιμάται ότι ικανοποιεί τις διάφορες ανάγκες που συνδέονται με τη συνήθη χρήση του κτιρίου (περιλαμβάνουν θέρμανση, ψύξη, εξαερισμό, φωτισμό, παραγωγή θερμού νερού κ.α.)
- Ενεργειακή κλάση κτιρίου: Η κατηγορία στην οποία κατατάσσεται ένα κτίριο ανάλογα με την ενεργειακή του απόδοση.
- Φορτίο θέρμανσης (heating load): Η ποσότητα της θερμικής ενέργειας που απαιτεί ένας χώρος για τη διατήρηση άνετων συνθηκών διαμονής.
- Φορτίο ψύξης (cooling load): Η ποσότητα της θερμικής ενέργειας που πρέπει να απομακρυνθεί από ένα χώρο (ψύξη) για τη διατήρηση άνετων συνθηκών διαμονής.
- Συνολική καταναλισκόμενη πρωτογενής ενέργεια κτιρίου: Το άθροισμα των επιμέρους υπολογιζόμενων ενεργειακών καταναλώσεων ενός κτιρίου για Θέρμανση-Ψύξη-Κλιματισμό (ΘΨΚ) και παραγωγής Ζεστού Νερού Χρήσης (ZNX). Σχετίζεται άμεσα με το φορτίο θέρμανσης και ψύξης, όπως θα αναλυθεί παρακάτω.

### 7.1.2 Συνοπτική παρουσίαση του προβλήματος

Στα κεφάλαια που ακολουθούν παρουσιάζεται αρχικά (Κεφάλαιο 7.2) το dataset των δεδομένων, που χρησιμοποιήθηκε στο πλαίσιο της παρούσας εργασίας για την τεκμηρίωση του προβλήματος ταξινόμησης που μελετάται και το οποίο διατίθεται από τους ερευνητές στην επιστημονικά έγκυρη βάση «Machine Learning Repository» του Πανεπιστημίου της California. Το dataset αυτό περιλαμβάνει τα αποτελέσματα των μετρήσεων για τα φορτία θέρμανσης και ψύξης, που έγιναν σε διαφορετικά κτίρια από την ομάδα των ερευνητών. Ο υπολογισμός των φορτίων θέρμανσης και ψύξης, πραγματοποιήθηκε με τη χρήση στατιστικών μεθόδων, βάσει 8 μεταβλητών/μεγεθών που παρουσιάζονται στο ίδιο κεφάλαιο. Παρατίθεται δείγμα του dataset, το οποίο αποτελεί παράρτημα και συνοδεύει την παρούσα εργασία.

Στη συνέχεια (Κεφάλαιο 7.3) γίνεται εκτενής αναφορά στη μεθοδολογία κατάταξης των κτιρίων σε ενεργειακές κλάσεις ανάλογα με τη μέση καταναλισκόμενη πρωτογενή ενέργεια κάθε κτιρίου και πώς αυτή υπολογίζεται από τα φορτία θέρμανσης και ψύξης,

στοιχεία που αντλούνται από το dataset του Πανεπιστημίου της California . Παρουσιάζονται τα διαδοχικά βήματα της μεθοδολογίας :

- Υπολογισμός συνολικής (ετήσιας) καταναλισκόμενης πρωτογενούς ενέργειας με τη χρήση των βαθμομερών θέρμανσης και ψύξης και με τη χρήση του φορτίου θέρμανσης και ψύξης.
- Υπολογισμός μέσης καταναλισκόμενης πρωτογενούς ενέργειας με τη χρήση της συνολικής (ετήσιας) καταναλισκόμενης πρωτογενούς ενέργειας και το σύνολο των βαθμομερών θέρμανσης και ψύξης.
- Σύγκριση της μέσης καταναλισκόμενης πρωτογενούς ενέργειας με την μέση καταναλισκόμενη ενέργεια του κτιρίου αναφοράς .
- Καθορισμός των ενεργειακών κλάσεων βάσει της σχέσης της μέσης ενεργειακής κατανάλωσης εκάστοτε κτιρίου με την αντίστοιχη του κτιρίου αναφοράς

Στη συνέχεια (Κεφάλαιο 8) υπάρχει η περιγραφή των πειραμάτων με το dataset το οποίο προκύπτει από τη μεθοδολογία που ακολουθήθηκε στο κεφάλαιο 7.3 ( το οποίο ουσιαστικά έχει τις ίδιες μετρήσεις των κτιρίων X1-X8 με το dataset του UCI και στη θέση των στηλών φορτίων θέρμανσης και ψύξης έχει την ενεργειακή κλάση του κάθε κτιρίου, όπως αυτή προκύπτει από τη μέση καταναλισκόμενη πρωτογενή ενέργεια που υπολογίστηκε στο κεφάλαιο 7.3). Η μελέτη για να θεωρηθεί πιο ολοκληρωμένη γίνεται για τρεις διαφορετικές πόλεις και πραγματοποιούνται πείραμα εκπαίδευσης του ELM και άλλων έξι αλγορίθμων.

Τέλος στο κεφάλαιο 9 συνοψίζονται παρατηρήσεις και συμπεράσματα που βασίζονται στα αποτελέσματα των πειραμάτων του κεφαλαίου 8 και αφορούν στην ακρίβεια της ταξινόμησης και στους χρόνους εκπαίδευσης.

## 7.2 Dataset (από UCI)

Ύστερα από ενδελεχή αναζήτηση, επιλέχθηκαν οι μετρήσεις και τα αποτελέσματα που παρουσιάζονται στο Machine Learning Repository του UCI (University of California). Τα αποτελέσματα αυτά αποτυπώνονται σε ένα dataset, που επισυνάπτεται στην παρούσα (<http://archive.ics.uci.edu/ml/datasets/Energy+efficiency>), δείγμα του οποίου εμφανίζεται παρακάτω :

A	B	C	D	E	F	G	H	I	J
X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2
0.98	514.50	294.00	110.25	7.00	2	0.00	0	15.55	21.33
0.98	514.50	294.00	110.25	7.00	3	0.00	0	15.55	21.33
0.98	514.50	294.00	110.25	7.00	4	0.00	0	15.55	21.33
0.98	514.50	294.00	110.25	7.00	5	0.00	0	15.55	21.33
0.90	563.50	318.50	122.50	7.00	2	0.00	0	20.84	28.28

Σχήμα 1: Δείγμα του dataset από το UCI

Ουσιαστικά, έχει αναπτυχθεί ένα στατιστικό πλαίσιο μηχανικής μάθησης, με σκοπό να μελετήσει την επίδραση οκτώ μεταβλητών εισόδου για τον υπολογισμό δύο μεταβλητών εξόδου. Συγκεκριμένα οι μεταβλητές εισόδου είναι :

- X1 : πόσο συμπαγές είναι το κτίριο
- X2 : ποια είναι η συνολική επιφάνειά του κτιρίου
- X3 : ποια είναι η συνολική επιφάνεια των τοίχων του κτιρίου
- X4: η συνολική επιφάνεια στέγης του κτιρίου
- X5 : το συνολικό ύψος του κτιρίου
- X6 : ο προσανατολισμός του κτιρίου
- X7 : η επιφάνεια των τζαμιών του κτιρίου και
- X8 : η κατανομή της επιφάνειας των τζαμιών του κτιρίου

Ενώ οι δύο μεταβλητές εξόδου είναι

- Το φορτίο θέρμανσης heating load (Y1) και
- Το φορτίο ψύξης cooling load (Y2).

Έχει ερευνηθεί συστηματικά η δύναμη σύνδεσης της κάθε μεταβλητής εισόδου με κάθε μια από τις μεταβλητές εξόδου, χρησιμοποιώντας μία ποικιλία εργαλείων από κλασσική και μη παραμετρική στατιστική ανάλυση, ώστε να εντοπιστούν οι «πιο στενά συνδεδεμένες» μεταβλητές εισόδου.

Το πρώτο βήμα, στις περισσότερες εφαρμογές ανάλυσης δεδομένων, είναι η μελέτη των στατιστικών ιδιοτήτων των μεταβλητών. Συνήθως αυτό επιτυγχάνεται με τη γραφική αναπαράσταση των πυκνοτήτων πιθανότητας, η οποία συνοψίζει την κάθε μεταβλητή για οπτικοποίηση. Ένας τρόπος να ληφθεί μία εμπειρική εκτίμηση της πυκνότητας είναι

τα ιστογράμματα. Αν και τα ιστογράμματα θεωρούνται αργά για απαιτητικές και πολύπλοκες αναλύσεις δεδομένων, έχουν το πλεονέκτημα του ότι δεν κάνουν εκτιμήσεις εκ των προτέρων για την κατανομή της εξεταζόμενης μεταβλητής και επιπλέον είναι πολύ εύκολος ο υπολογισμός τους. Συχνά αυτό το στάδιο υπολογισμού αποκαλύπτει αν η μεταβλητή ακολουθεί Gauss -ιανή κατανομή και συνεπώς χαρακτηρίζεται από μία κορυφή στη μέση του δυνατού φάσματος τιμών και από συμμετρία.

Στο συγκεκριμένο dataset τα δεδομένα δεν είναι Gauss -ιανά για αυτό και χρησιμοποιείται ο συντελεστής συσχέτισης Spearman, για να μπορέσει να βγει κάποιο συμπέρασμα σχετικά με το πόσο στενά συνδεδεμένα είναι η ισχύς της κάθε μεταβλητής εισόδου με καθεμιά έξοδο. Η αξιολόγηση αυτής της σύνδεσης γίνεται με τον έλεγχο των σφαλμάτων  $p$  ( $p$ -value). Επιπλέον ελέγχεται η σημαντικότητα σε επίπεδο 0.01.

Το dataset αποτελείται από 768 δείγματα και 8 μεταβλητές εισόδου. Είναι εύκολο να κατασκευαστεί η μήτρα  $X \in \mathbb{R}^{768 \times 8}$  και στη συνέχεια πρέπει να βρεθεί μία σχέση μεταξύ της μήτρας και των αποκρίσεων εξόδου  $Y_1$  (*heating load*) και  $Y_2$  (*cooling load*), τέτοια ώστε  $Y=f(X)$ . Για να προσεγγιστεί το πρόβλημα χρησιμοποιούνται δύο «τρόποι εκπαίδευσης». Ο πρώτος έγκειται σε μία ισχυρή τυποποιημένη τεχνική παλινδρόμησης (επαναληπτική και επανασταθμιζόμενη μέθοδος ελαχίστων τετραγώνων) και ο δεύτερος σε έναν ισχυρό, μη γραμμικό και μη παραμετρικό ταξινομητή. Αφού τελειώσει η εκπαίδευση, ελέγχεται η απόδοση γενίκευσης με τη μέθοδο cross validation και στη συνέχεια καταγράφεται το μέσο απόλυτο σφάλμα, το μέσο τετραγωνικό σφάλμα και το μέσο σχετικό σφάλμα.

Η μελέτη, όπως προαναφέρθηκε, περιλαμβάνει 768 κτίρια κατοικιών, τα οποία έχουν τον ίδιο όγκο αλλά διαφοροποιούνται στις επιφάνειες και στις διαστάσεις τους. Τα κατασκευαστικά υλικά είναι κοινά για όλα τα κτίρια. Επιπλέον χρησιμοποιούνται τρία διαφορετικά ποσοστά επιφάνειας γυαλιού προς τη συνολική επιφάνεια (10%, 25% και 40%) και πέντε διαφορετικές κατανομές της συνολικής επιφάνειας του γυαλιού στο κτίριο (ομοιόμορφη - 25% τζάμια σε κάθε πλευρά, βόρεια - 55% γυαλί στη βόρεια πλευρά και 15% στις υπόλοιπες, ανατολική - 55% γυαλί στην ανατολική πλευρά και 15% στις υπόλοιπες, νότια - 55% γυαλί στη νότια πλευρά και 15% στις υπόλοιπες και δυτική - 55% γυαλί στη δυτική πλευρά και 15% στις υπόλοιπες).

### 7.3 Πρόβλημα ταξινόμησης ενεργειακών κλάσεων

Στο παραπάνω dataset όλα όσα προαναφέρθηκαν έχουν συνυπολογιστεί για τον υπολογισμό φορτίου θέρμανσης και ψύξης. Όταν πρόκειται για τον σχεδιασμό ενεργειακά αποτελεσματικών κτιρίων, απαιτείται ο υπολογισμός του φορτίου θέρμανσης (*heating load*) και του φορτίου ψύξης (*cooling load*) για να μπορέσουν να αποτυπωθούν οι προδιαγραφές του εξοπλισμού θέρμανσης και ψύξης που απαιτούνται για τη διατήρηση άνετων συνθηκών διαμονής.

Η ταξινόμηση, στο υπό μελέτη πρόβλημα ταξινόμησης, αφορά στην κατάταξη των κτιρίων βάσει της ενεργειακής τους κλάσης. Η ενεργειακή τους κλάση καθορίζεται από τη μέση καταναλισκόμενη πρωτογενή ενέργεια.

Η συνολική (ετήσια) καταναλισκόμενη πρωτογενής ενέργεια, προκύπτει με την χρήση των  $Y_1$  και  $Y_2$ , υπολογίζεται ως το άθροισμα του γινομένου των βαθμομερών θέρμανσης επί το heating load ( $Y_1$ ) και του γινομένου των βαθμομερών ψύξης επί το cooling load ( $Y_2$ ) και στη συνέχεια υπολογίζεται η μέση κατανάλωση ενέργεια (το προηγούμενο άθροισμα διὰ το σύνολο όλων των βαθμομερών θέρμανσης και ψύξης), ο οποίος συγκρίνεται με τον αντίστοιχο που έχει θεωρηθεί για το κτίριο αναφοράς.

### **7.3.1 Κτίριο Αναφοράς**

Η κατάταξη του εκάστοτε εξεταζόμενου κτιρίου σε μία από τις ενεργειακές κλάσεις έγινε μέσω σύγκρισης της μέσης κατανάλωσης πρωτογενούς ενέργειας του εξεταζόμενου κτιρίου, σε σχέση με την μέση κατανάλωση του κτιρίου αναφοράς. Το κτίριο αναφοράς είναι ένα κτίριο με το ίδιο προφίλ και με ίδιες συνθήκες λειτουργίας με τα προς μελέτη κτίρια. Συγκεκριμένα θεωρείται ίδιο σε γεωμετρικά χαρακτηριστικά, θέση, προσανατολισμό, χρήση και χαρακτηριστικά λειτουργίας με το προς εξέταση κτίριο. Το κτίριο αναφοράς πληροί τις ελάχιστες απαιτήσεις θερμομόνωσης κτιριακού κελύφους και ηλεκτρομηχανολογικών εγκαταστάσεων που αφορούν στην θέρμανση και ψύξη των εσωτερικών χώρων, στην παραγωγή ζεστού νερού χρήσης και στο φωτισμό.

### **7.3.2 Βαθμομέρες και βαθμώρες**

Οι βαθμομέρες χρησιμεύουν στην εκτίμηση των ενεργειακών αναγκών για τη θέρμανση ή ψύξη ενός κτιρίου. Είναι μία μέθοδος, η οποία θεωρείται από τις πιο διαδεδομένες και απλές μεθόδους για την εκτίμηση της ενεργειακής κατανάλωσης τόσο για τη θέρμανση όσο και για τη ψύξη των κτιρίων. Στη διεθνή βιβλιογραφία έχουν προταθεί και διάφορες άλλες μέθοδοι για τον υπολογισμό των μεγεθών αυτών, οι περισσότερες εκ των οποίων απαιτούν τη γνώση αναλυτικών θερμοκρασιακών δεδομένων. Συνήθως η πρόσβαση σε αναλυτικά μετεωρολογικά δεδομένα δεν είναι δυνατή, με αποτέλεσμα η χρήση των αναλυτικών μεθόδων να αποτελεί μια δύσκολη ή και πολλές φορές αδύνατη διαδικασία.

#### **7.3.2.1 Βαθμομέρες θέρμανσης**

Η παράμετρος αυτή προκύπτει από τη μέση μηνιαία θερμοκρασία αέρα εικοσιτετράωρου για τους θερινούς μήνες σε μία περιοχή και μία θερμοκρασία

αναφοράς η οποία συνήθως λαμβάνεται ίση με 18°C. Η εξίσωση υπολογισμού των βαθμομερών θέρμανσης DD είναι:

$$DD = \sum [N_{mo} * (T_{av} - T_a) +] \quad (35)$$

Όπου,  $N_{mo}$  ο αριθμός ημερών για κάθε μήνα,  $T_{av}$  η θερμοκρασία αναφοράς (°C) και  $T_a$  η μηνιαία θερμοκρασία αέρα περιβάλλοντος (°C). Το σύμβολο (+) δηλώνει ότι στους υπολογισμούς λαμβάνονται υπόψη μόνο οι περιπτώσεις που η θερμοκρασία  $T_a$  δεν υπερβαίνει τη θερμοκρασία αναφοράς, δηλαδή έχουμε θετική τιμή στην παρένθεση)

### 7.3.2.2 Βαθμώρες ψύξης

Για τον υπολογισμό των απαιτούμενων φορτίων ψύξης ενός κτιρίου, χρησιμοποιούνται οι βαθμώρες ψύξης. Η παράμετρος αυτή προκύπτει από την μέση ωριαία θερμοκρασία τους θερινούς μήνες σε μία περιοχή και μία θερμοκρασία αναφοράς η οποία συνήθως λαμβάνεται ίση με 26 °C. Η εξίσωση υπολογισμού των βαθμωρών ψύξης CDH είναι:

$$CDH = \sum [N_{mo} * (T_{ai} - T_{av}) +] \quad (36)$$

Όπου,  $N_{mo}$  ο αριθμός ημερών για κάθε μήνα,  $T_{av}$  η θερμοκρασία αναφοράς (°C) και  $T_{ai}$  η μέση ωριαία θερμοκρασία αέρα περιβάλλοντος (°C). Το σύμβολο (+) δηλώνει ότι στους υπολογισμούς λαμβάνονται υπόψη μόνο οι περιπτώσεις που η θερμοκρασία  $T_{ai}$  υπερβαίνει τη θερμοκρασία αναφοράς, δηλαδή έχουμε θετική τιμή στην παρένθεση).

Για να θεωρηθεί η μελέτη μας ακόμα πιο ακριβής, πρέπει να εξεταστεί κατά πόσο επηρεάζει η γεωγραφική θέση του προς εξέταση κτιρίου, και θεωρώντας ότι έχουμε την ίδια κατανάλωση πρωτογενούς ενέργειας για θέρμανση και ψύξη εξετάστηκαν τρεις διαφορετικές περιπτώσεις σε τρεις διαφορετικές πόλεις (Αθήνα, Θεσσαλονίκη, Καστοριά), οι οποίες ανήκουν σε διαφορετικές κλιματικές ζώνες.

Στους δύο παρακάτω πίνακες φαίνονται οι βαθμομέρες θέρμανσης για διαφορετικές περιοχές της Ελλάδος, καθώς και οι βαθμώρες ψύξης:



Πίνακας 3.7. Βαθμομέρες θέρμανσης DD με θερμοκρασία αναφοράς 18°C

Περιοχή/μήνας	ΙΑΝ	ΦΕΒ	ΜΑΡ	ΑΠΡ	ΜΑΙ	ΙΟΥΝ	ΙΟΥΛ	ΑΥΓ	ΣΕΠ	ΟΚΤ	ΝΟΕ	ΔΕΚ
Αθήνα (Ελληνικό)	239	207	177	60	-	-	-	-	-	-	78	186
Αθήνα (Φιλαδέλφεια)	239	207	177	60	-	-	-	-	-	-	78	186
Αγρίνιο	301	246	202	84	-	-	-	-	-	-	144	260
Αγχίαλος	353	288	248	111	-	-	-	-	-	31	177	307
Αλεξανδρούπολη	403	339	301	144	-	-	-	-	-	71	216	341
Αλιάρτος	338	274	229	84	-	-	-	-	-	34	180	291
Ανδραβίδα	267	227	192	96	-	-	-	-	-	-	117	220
Άραξος	242	210	180	84	-	-	-	-	-	-	99	198
Άργος (Πυργέλα)	307	269	229	99	-	-	-	-	-	6	153	264
Αργοστόλι	205	182	161	81	-	-	-	-	-	-	66	161
Άρτα	310	252	202	102	-	-	-	-	-	6	165	276
Δράμα	412	322	248	99	-	-	-	-	-	50	243	372
Έδεσσα	419	330	260	138	-	-	-	-	-	74	258	378
Ζάκυνθος	233	210	192	96	-	-	-	-	-	-	96	195
Ηράκλειο	183	162	140	45	-	-	-	-	-	-	39	133
Θεσσαλονίκη	394	314	254	111	-	-	-	-	-	53	207	344
Ιεράπετρα	158	143	118	30	-	-	-	-	-	-	15	109
Ιωάννινα	412	336	285	168	16	-	-	-	-	93	249	378
Καλαμάτα	242	207	177	84	-	-	-	-	-	-	96	198
Καρδίτσα	419	311	236	123	-	-	-	-	-	59	237	425
Καρπενήσι	440	417	391	222	102	-	-	-	-	174	336	409
Κάρυστος	236	216	174	69	-	-	-	-	-	-	93	195
Καστοριά	490	409	344	195	50	-	-	-	-	143	324	465
Κέρκυρα	257	216	186	90	-	-	-	-	-	-	111	214
Κοζάνη	487	400	344	192	37	-	-	-	-	140	300	437
Κομοτηνή	409	330	291	147	-	-	-	-	-	87	216	341
Κόνιτσα	397	322	264	174	25	-	-	-	-	81	246	380
Κόρινθος (Βέλο)	285	244	202	78	-	-	-	-	-	-	138	245
Κύθηρα	220	199	189	102	-	-	-	-	-	-	66	164
Κως	217	210	183	78	-	-	-	-	-	-	90	174
Λαμία	338	277	226	90	-	-	-	-	-	31	183	304
Λάρισα	397	314	264	120	-	-	-	-	-	53	213	357
Λευκάδα	242	207	164	84	-	-	-	-	-	-	102	202
Λήμνος	329	286	257	126	-	-	-	-	-	37	171	279
Μεθώνη	208	182	158	78	-	-	-	-	-	-	60	158
Μήλος	226	202	189	90	-	-	-	-	-	-	81	174
Μυτιλήνη	264	227	198	72	-	-	-	-	-	-	111	208
Νάξος	183	162	146	57	-	-	-	-	-	-	51	136
Ξάνθη	384	314	260	111	-	-	-	-	-	47	210	344
Πάρος	211	190	158	54	-	-	-	-	-	-	84	177
Πάτρα	248	207	171	72	-	-	-	-	-	-	105	205
Πολύγυρος	406	372	288	168	53	-	-	-	-	90	219	347
Πύργος	260	221	180	84	-	-	-	-	-	-	117	217
Ρέθυμνο	161	143	118	27	-	-	-	-	-	-	21	109
Ρόδος	186	162	133	42	-	-	-	-	-	-	39	140
Σάμος	236	218	180	57	-	-	-	-	-	-	99	186
Σέρρες	434	328	257	108	-	-	-	-	-	71	258	409
Σητεία	180	160	136	42	-	-	-	-	-	-	36	130
Σκύρος	251	218	202	90	-	-	-	-	-	-	102	198
Σούδα	220	196	167	66	-	-	-	-	-	-	72	167

Πίνακας 1: Βαθμομέρες θέρμανσης DD με θερμοκρασία αναφοράς 18 °C

**Πίνακας 3.8.** Βαθμώρες ψύξης CDH με θερμοκρασία αναφοράς 26°C

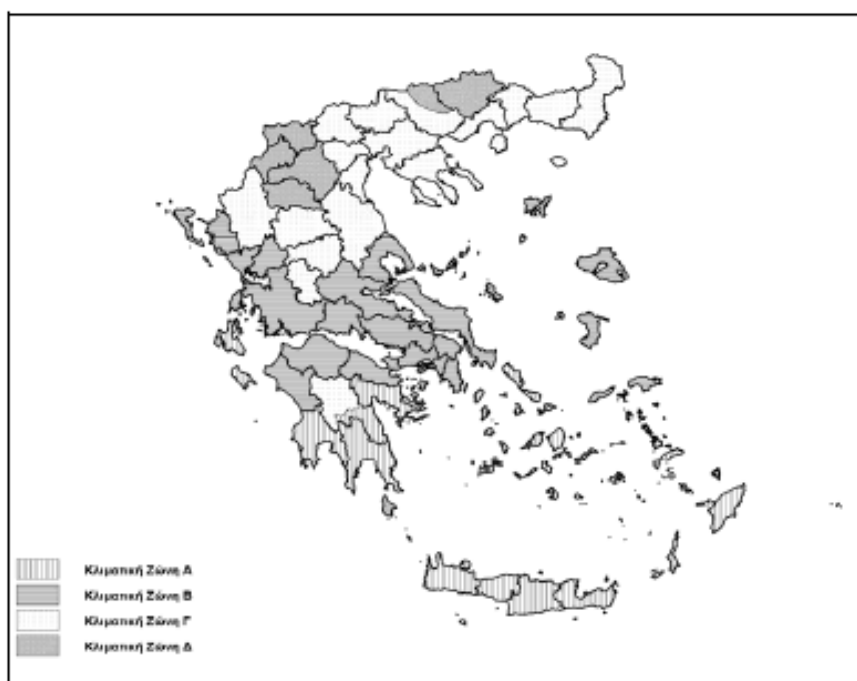
Περιοχή/μήνας	ΙΑΝ	ΦΕΒ	ΜΑΡ	ΑΠΡ	ΜΑΙ	ΙΟΥΝ	ΙΟΥΛ	ΑΥΓ	ΣΕΠ	ΟΚΤ	ΝΟΕ	ΔΕΚ
Αθήνα (Ελληνικό)	-	-	-	-	-	794	1901	1853	292	-	-	-
Αθήνα (Φιλαδέλφεια)	-	-	-	-	13	1209	2162	1787	363	-	-	-
Αγρίνιο	-	-	-	-	16	909	1535	1488	248	-	-	-
Αγχιάλος	-	-	-	-	-	552	1281	1042	34	-	-	-
Αλεξανδρούπολη	-	-	-	-	-	152	917	808	-	-	-	-
Ανδραβίδα	-	-	-	-	-	324	799	1065	109	-	-	-
Άργος (Πυργέλα)	-	-	-	-	-	945	1813	1572	372	-	-	-
Αργοστόλι	-	-	-	-	-	59	626	873	-	-	-	-
Άρτα	-	-	-	-	-	675	1290	1297	137	-	-	-
Ηράκλειο	-	-	-	-	-	12	694	610	18	-	-	-
Θεσσαλονίκη	-	-	-	-	-	526	1211	1058	-	-	-	-
Ιωάννινα	-	-	-	-	-	219	800	675	-	-	-	-
Καλαμάτα	-	-	-	-	-	461	1146	1130	187	-	-	-
Καστοριά	-	-	-	-	-	55	505	369	-	-	-	-
Κέρκυρα	-	-	-	-	-	391	1122	1236	1	-	-	-
Κόρινθος (Βέλο)	-	-	-	-	-	844	1698	1560	236	-	-	-
Λαμία	-	-	-	-	-	921	1496	1181	181	-	-	-
Λάρισα	-	-	-	-	-	951	1563	1221	121	-	-	-
Νάξος	-	-	-	-	-	-	166	100	-	-	-	-
Πύργος	-	-	-	-	-	486	1246	1440	217	-	-	-
Ρόδος	-	-	-	-	-	158	870	1046	161	-	-	-
Σάμος	-	-	-	-	-	771	1867	1782	212	-	-	-
Σέρρες	-	-	-	-	-	688	1344	1158	59	-	-	-
Σητεία	-	-	-	-	-	-	408	350	-	-	-	-
Σκύρος	-	-	-	-	-	-	363	218	-	-	-	-
Σούδα	-	-	-	-	-	497	1276	1051	157	-	-	-
Τανάγρα	-	-	-	-	-	753	1557	1235	128	-	-	-
Τρίκαλα Ημαθ.	-	-	-	-	-	591	1027	830	9	-	-	-
Τυμπάκι	-	-	-	-	-	376	1476	1512	267	-	-	-
Χρυσούπολη	-	-	-	-	-	49	674	559	-	-	-	-
Λαμία	-	-	-	-	-	794	1901	1853	292	-	-	-

Πίνακας 2: Βαθμώρες ψύξης CDH με θερμοκρασία αναφοράς 26°C

### 7.3.3 Κλιματικές ζώνες

Σύμφωνα με τον Κανονισμό Ενεργειακής Απόδοσης Κτιρίων (Κ.Εν.Α.Κ), η ελληνική επικράτεια διαιρείται σε τέσσερις κλιματικές ζώνες βάσει των βαθμομερών θέρμανσης. Σε κάθε νομό, οι περιοχές που βρίσκονται σε υψόμετρο άνω των 500 μέτρων, εντάσσονται στην επόμενη ψυχρότερη κλιματική ζώνη από εκείνη στην οποία ανήκουν σύμφωνα με την παρακάτω εικόνα (Εξαίρεση αποτελεί η ζώνη Δ διότι όλες οι περιοχές ανεξαρτήτως υψόμετρου περιλαμβάνονται στη ζώνη Δ)

ΚΛΙΜΑΤΙΚΗ ΖΩΝΗ	ΝΟΜΟΙ
ΖΩΝΗ Α	Ηρακλείου, Χανίων, Ρεθύμνου, Λασιθίου, Κυκλάδων, Δωδεκανήσου, Σάμου, Μεσσηνίας, Λακωνίας, Αργολίδας, Ζακύνθου, Κεφαλληνίας & Ιθάκης, Κύθηρα & νησιά Σαρωνικού (Αττικής), Αρκαδίας (πεδινή).
ΖΩΝΗ Β	Αττικής (εκτός Κυθέρων & νησιών Σαρωνικού), Κορινθίας, Ηλείας, Αχαΐας, Αιτωλοακαρνανίας, Φθιώτιδας, Φωκίδας, Βοιωτίας, Ευβοίας, Μαγνησίας, Λέσβου, Χίου, Κέρκυρας, Λευκάδας, Θεσπρωτίας, Πρέβεζας, Άρτας.
ΖΩΝΗ Γ	Αρκαδίας (ορεινή), Ευρυτανίας, Ιωαννίνων, Λάρισας, Καρδίτσας, Τρικάλων, Πιερίας, Ημαθίας, Πέλλας, Θεσσαλονίκης, Κιλκίς, Χαλκιδικής, Σερρών (εκτός ΒΑ τμήματος), Καβάλας, Ξάνθης, Ροδόπης, Έβρου.
ΖΩΝΗ Δ	Γρεβενών, Κοζάνης, Καστοριάς, Φλώρινας, Σερρών (ΒΑ τμήμα), Δράμας.



Εικόνα 3: Σχηματική απεικόνιση κλιματικών ζωνών ελληνικής επικράτειας

### 7.3.4 Ενεργειακές κλάσεις

Όπως προαναφέρθηκε, μέσω των βαθμομερών θέρμανσης και των βαθμοωρών ψύξης για την κάθε πόλη (Αθήνα, Θεσσαλονίκη, Καστοριά) και με δεδομένα τα φορτία θέρμανσης (*heating load*) και ψύξης (*cooling load*) από το dataset του UCI μπορεί να υπολογιστεί η συνολική (ετήσια) καταναλισκόμενη πρωτογενής ενέργεια και η μέση καταναλισκόμενη πρωτογενής ενέργεια για το κάθε κτίριο χωριστά. Στη συνέχεια με

βοήθεια του κτιρίου αναφοράς κατατάσσεται εύκολα στην ενεργειακή του κλάση, σύμφωνα με τον παρακάτω πίνακα:

Κατηγορία	Όρια κατηγορίας	Όρια κατηγορίας
A+	$EP \leq 0,33R_R$	$T \leq 0,33$
A	$0,33R_R < EP \leq 0,50R_R$	$0,33 < T \leq 0,50$
B+	$0,50R_R < EP \leq 0,75R_R$	$0,50 < T \leq 0,75$
B	$0,75R_R < EP \leq 1,00R_R$	$0,75 < T \leq 1,00$
Γ	$1,00R_R < EP \leq 1,41R_R$	$1,00 < T \leq 1,41$
Δ	$1,41R_R < EP \leq 1,82R_R$	$1,41 < T \leq 1,82$
E	$1,82R_R < EP \leq 2,27R_R$	$1,82 < T \leq 2,27$
Z	$2,27R_R < EP \leq 2,73R_R$	$2,27 < T \leq 2,73$
H	$2,73R_R < EP$	$2,73 < T$

Πίνακας 3: Κατηγορίες ενεργειακής απόδοσης κτιρίων

Όπου ο δείκτης  $R_R$  είναι ίσος με τη μέση υπολογιζόμενη κατανάλωση πρωτογενούς ενέργειας του κτιρίου αναφοράς. Ο λόγος  $T$  είναι το πηλίκο της μέσης υπολογιζόμενης κατανάλωσης πρωτογενούς ενέργειας του εξεταζόμενου κτηρίου (EP) προς την τιμή της υπολογιζόμενης μέσης κατανάλωσης πρωτογενούς ενέργειας του κτιρίου αναφοράς ( $R_R$ ) και αποτελεί το κριτήριο για την κατάταξη του κτηρίου στην αντίστοιχη κατηγορία ενεργειακής απόδοσης.

### 7.3.5 Παραδοχές

- Για τη συγκεκριμένη μελέτη θεωρείται ότι το προς εξέταση κτίριο είναι Κατοικία, χρησιμοποιείται ηλιακός θερμοσίφωνας για την κάλυψη των αναγκών ζεστού νερού χρήσης την θερινή περίοδο ενώ για την περίπτωση της χειμερινής περιόδου η καταναλισκόμενη ενέργεια για την παραγωγή ζεστού νερού χρήσης εμπεριέχεται στην ολικά αναγραφόμενη καταναλισκόμενη ενέργεια θέρμανσης.
- Ο φωτισμός γίνεται με φωτιστικά τύπου LED με πολύ μικρή εκπομπή θερμότητας. (Ειδικά για τα κτίρια κατοικίας στη συνολική ενεργειακή κατανάλωση δεν συμπεριλαμβάνεται ο φωτισμός).
- Το σύστημα θέρμανσης για όλες τις περιπτώσεις είναι ένας λέβητας πετρελαίου υψηλής απόδοσης ενώ οι τερματικές μονάδες είναι άμεσης απόδοσης.
- Το σύστημα ψύξης γίνεται με ηλεκτρικές μονάδες τύπου split

## 8. Πειράματα

Το τελικό dataset, βάσει της προαναφερθείσας μεθοδολογίας, έχει 9 στήλες. Οι πρώτες οχτώ στήλες είναι οι μεταβλητές (X1-X8, οι οποίες έχουν να κάνουν με τον προσανατολισμό του κτιρίου, την επιφάνεια της ταράτσας, των παραθύρων, του τοίχου που έρχεται σε επαφή με εξωτερικούς χώρους κτλ) και η ένατη στήλη είναι η ενεργειακή κλάση του κτιρίου, που προκύπτει από τις μεταβλητές Y1 και Y2 του αρχικού dataset (Y1=heating load, Y2 = cooling load).

Επιλέχθηκαν τρεις διαφορετικές πόλεις για την πραγματοποίηση των πειραμάτων (Ν. Φιλαδέλφεια, Θεσσαλονίκη και Καστοριά). Σύμφωνα με όσα προαναφέρθηκαν, οι βαθμομέρες θέρμανσης και ψύξης διαφοροποιούνται ανάλογα με την κάθε περιοχή και η κατανάλωση πρωτογενούς ενέργειας για κάθε κτίριο είναι διαφορετική ανάλογα με την πόλη που βρίσκεται, συνεπώς μπορεί να υπάρχει διαφοροποίηση στην ενεργειακή κλάση που ανήκει. Αυτό έχει σαν αποτέλεσμα να διαφοροποιείται η ενεργειακή κλάση του κτιρίου (ένατη στήλη του dataset) ανάλογα με την πόλη που βρίσκεται το κτίριο.

Με αυτόν τον τρόπο δημιουργήθηκαν 3 διαφορετικά dataset, ένα για την κάθε πόλη). Στο αρχείο dedomena\_all.xlsx φαίνονται αναλυτικότερα οι υπολογισμοί και τα αποτελέσματά τους. Ενδεικτικά παρατίθεται απόσπασμα του συγκεκριμένου excel :

A	B	C	D	E	F	G	H	I	J	K	L
Ν.ΦΙΛΑΔΕΛΦΕΙΑ (ΑΘΗΝΑ)						ΘΕΣΣΑΛΟΝΙΚΗ					
HL(Y1)	CL(Y2)	ΣΥΝΟΛΙΚΗ (ΕΤΗΣΙΑ) ΚΑΤΑΝΑΛΙΣΚΟΜΕΝΗ ΠΡΩΤΟΓΕΝΗΣ ΕΝΕΡΓΕΙΑ	ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΣΗΣ ΚΑΤΑΝΑΛΙΣΚΟΜΕΝΗΣ Σ ΠΡΩΤΟΓΕΝΟΥΣ ΕΝΕΡΓΕΙΑΣ	ΕΝΕΡΓΕΙΑΚΗ ΚΛΑΣΗnumber	Energy_Class	ΣΥΝΟΛΙΚΗ (ΕΤΗΣΙΑ) ΚΑΤΑΝΑΛΙΣΚΟΜΕΝΗ ΠΡΩΤΟΓΕΝΗΣ ΕΝΕΡΓΕΙΑ	ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΣΗΣ ΚΑΤΑΝΑΛΙΣΚΟΜΕΝΗΣ ΠΡΩΤΟΓΕΝΟΥΣ ΕΝΕΡΓΕΙΑΣ	ΕΝΕΡΓΕΙΑΚΗ ΚΛΑΣΗnumber	Energy_Class	ΣΥΝΟΛΙΚΗ (ΕΤΗΣΙΑ) ΚΑΤΑΝΑΛΙΣΚΟΜΕΝΗ ΠΡΩΤΟΓΕΝΗΣ ΕΝΕΡΓΕΙΑ	
		15.55	21.33	19653.08	16.68342954	4	C	28572.96	15.92695652	3	B
15.55	21.33	19653.08	16.68342954	4	C	28572.96	15.92695652	3	B	38462.87	
15.55	21.33	19653.08	16.68342954	4	C	28572.96	15.92695652	3	B	38462.87	
15.55	21.33	19653.08	16.68342954	4	C	28572.96	15.92695652	3	B	38462.87	
20.84	28.28	26268.16	22.29894737	4	C	38257.44	21.32521739	4	C	51535.72	
21.46	25.38	26185.4	22.2286927	4	C	38957.88	21.71565217	4	C	52923.02	
20.71	25.16	25424.33	21.58262309	4	C	37674.39	21.00021739	4	C	51099.44	
19.68	29.60	25474.56	21.62526316	4	C	36466.56	20.32695652	4	C	48780	

Σχήμα 2: Δείγμα του αρχείου excel με τους αναλυτικούς υπολογισμούς για τις ενεργειακές κλάσεις (μέρος 1)

L	M	N	O	P	Q	R	S	T	U	V	W
<b>ΚΑΣΤΟΡΙΑ</b>											
ΣΥΝΟΛΙΚΗ (ΕΤΗΣΙΑ) ΚΑΤΑΝΑΛΙΣΚΟΜΕΝΗ ΠΡΩΤΟΓΕΝΗΣ ΕΝΕΡΓΕΙΑ	ΣΥΝΤΕΛΕΣΤΗΣ (ΜΕΣΗΣ) ΚΑΤΑΝΑΛΙΣΚΟΜΕΝΗ Σ ΠΡΩΤΟΓΕΝΟΥΣ ΕΝΕΡΓΕΙΑΣ	ΕΝΕΡΓΕΙΑΚΗ ΚΛΑΣΗnumber	Energy_Class	X1	X2	X3	X4	X5	X6	X7	X8
38462.87	15.64167141	3	B	0.98	514.50	294.00	110.25	7.00	2	0.00	0
38462.87	15.64167141	3	B	0.98	514.50	294.00	110.25	7.00	3	0.00	0
38462.87	15.64167141	3	B	0.98	514.50	294.00	110.25	7.00	4	0.00	0
38462.87	15.64167141	3	B	0.98	514.50	294.00	110.25	7.00	5	0.00	0
51535.72	20.95799919	4	C	0.90	563.50	318.50	122.50	7.00	2	0.00	0
52923.02	21.52217161	4	C	0.90	563.50	318.50	122.50	7.00	3	0.00	0
51099.44	20.78057747	4	C	0.90	563.50	318.50	122.50	7.00	4	0.00	0
48780	19.83733225	4	C	0.90	563.50	318.50	122.50	7.00	5	0.00	0

Σχήμα 3: Δείγμα του αρχείου excel με τους αναλυτικούς υπολογισμούς για τις ενεργειακές κλάσεις (μέρος 2)

A	B	C	D	E	F
<b>ΝΕΑ ΦΙΛΑΔΕΛΦΕΙΑ (ΑΘΗΝΑ)</b>		<b>ΘΕΣΣΑΛΟΝΙΚΗ</b>		<b>ΚΑΣΤΟΡΙΑ</b>	
ΒΑΘΜΟΗΜΕΡΕΣ ΘΕΡΜΑΝΣΗΣ=947	ΒΑΘΜΟΗΜΕΡΕΣ ΨΥΞΗΣ=231	ΒΑΘΜΟΗΜΕΡΕΣ ΘΕΡΜΑΝΣΗΣ=1677	ΒΑΘΜΟΗΜΕΡΕΣ ΨΥΞΗΣ=117	ΒΑΘΜΟΗΜΕΡΕΣ ΘΕΡΜΑΝΣΗΣ=1677	ΒΑΘΜΟΗΜΕΡΕΣ ΨΥΞΗΣ=117

Σχήμα 4: Βαθμοημέρες θέρμανσης και ψύξης για τις τρεις πόλεις των πειραμάτων

Στο σχήμα 4 αποτυπώνονται οι βαθμοημέρες θέρμανσης και ψύξης για τις τρεις πόλεις που επιλέχθηκαν για τα πειράματα και βάσει των οποίων υπολογίζονται οι τιμές της συνολικής καταναλισκόμενης πρωτογενούς ενέργειας και του συντελεστή μέσης συνολικής καταναλισκόμενης πρωτογενούς ενέργειας στα Σχήματα 2 και 3.

Σημειώνεται, ότι οι βαθμοημέρες θέρμανσης προέκυψαν από τον πίνακα 1 αθροίζοντας τις ημέρες από κάθε μήνα για την ίδια πόλη (π.χ. για την Ν. Φιλαδέλφεια είναι  $239+207+177+60+78+186=947$ ) ενώ οι βαθμοημέρες ψύξης προέκυψαν από τον πίνακα 2 με τον ίδιο τρόπο αλλά διαιρώντας το τελικό άθροισμα δια του 24 καθώς ο πίνακας αναφέρεται σε βαθμώρες (π.χ. για την Ν. Φιλαδέλφεια  $(13+1209+2162+1787+363)/24=231$ ).

Η πηγή άντλησης των συγκεκριμένων τιμών είναι η τεχνική οδηγία του ΤΕΕ (Τ.Ο.Τ.Ε.Ε. 20701-1/2010 – Αναλυτικές εθνικές προδιαγραφές παραμέτρων για τον υπολογισμό της ενεργειακής απόδοσης κτιρίων και την έκδοση του πιστοποιητικού ενεργειακής απόδοσης).

Στα σχήματα 2 και 3 εμφανίζονται :

Οι δύο μεταβλητές εξόδου (Το φορτίο θέρμανσης heating load (Y1) και το φορτίο ψύξης cooling load (Y2)) που προκύπτουν βάσει αλγορίθμου από τις μεταβλητές X1 έως X8.

Η συνολική (ετήσια) καταναλισκόμενη πρωτογενής ενέργεια (Στήλη C), η οποία προκύπτει από το άθροισμα των γινομένων της στήλης A (δηλαδή το φορτίο θέρμανσης) επί τον αριθμό 947 (οι βαθμοήμερες θέρμανσης, που χαρακτηρίζουν τη Νέα Φιλαδέλφεια) συν τη στήλη B (δηλαδή το φορτίο ψύξης) επί τον αριθμό 231 (βαθμοήμερες ψύξης, ομοίως για τη Ν. Φιλαδέλφεια).

Ο συντελεστής της μέσης καταναλισκόμενης πρωτογενούς ενέργειας (Στήλη D), ο οποίος προκύπτει από το πηλίκο της στήλης C διά τον συνολικό αριθμό βαθμοημερών θέρμανσης και ψύξης ( $941+231=1172$  βαθμοήμερες συνολικά για την περίπτωση της Νέας Φιλαδέλφειας).

Η ανάγκη εισαγωγής στο πείραμα μας, του συντελεστή της μέσης καταναλισκόμενης πρωτογενούς ενέργειας, τεκμηριώνεται από την απουσία μονάδων μέτρησης των Y1 (φορτίο θέρμανσης) και Y2 (φορτίο ψύξης) και συνεπώς την αδυναμία ακριβούς προσδιορισμού της συνολικής καταναλισκόμενης πρωτογενής ενέργειας.

Παρατηρώντας τη στήλη D, διαπιστώνεται ότι οι τιμές του συντελεστή μέσης καταναλισκόμενης πρωτογενούς ενέργειας κυμαίνονται στο εύρος τιμών 7 έως 43.

Για ένα κτίριο με χαμηλό συντελεστή μέσης καταναλισκόμενης πρωτογενούς ενέργειας (στην περίπτωση μας κοντά στο 7) ισχύει το εξής:

Το συγκεκριμένο κτίριο έχει ελάχιστες απαιτήσεις για επιπλέον εγκαταστάσεις θέρμανσης και ψύξης, συνεπώς είναι αρκετά αποδοτικό αφού δεν απαιτούνται μεγάλα ποσά ενέργειας για να εξασφαλίσει τη διατήρηση άνετων συνθηκών διαμονής.

Αντίθετα, για ένα κτίριο με υψηλό συντελεστή μέσης καταναλισκόμενης πρωτογενούς ενέργειας (στην περίπτωση μας κοντά στο 43) ισχύει το εξής:

Το συγκεκριμένο κτίριο έχει μέγιστες απαιτήσεις για επιπλέον εγκαταστάσεις θέρμανσης και ψύξης, συνεπώς δεν είναι καθόλου αποδοτικό αφού απαιτούνται μεγάλα ποσά ενέργειας για να εξασφαλίσει τη διατήρηση άνετων συνθηκών διαμονής.

Στη συγκεκριμένη μελέτη μια καλής προσέγγισης κλίμακα θεωρείται η κλίμακα με εύρος τιμών από 0 έως 50 και όχι με εύρος τιμών από 7 έως 43 για τον παρακάτω λόγο:

Δεν υπάρχουν κτίρια τα οποία να είναι τόσο μη αποδοτικά ενεργειακά ώστε να ακουμπούν το άνω όριο του εύρους τιμών (για παράδειγμα δεν είναι δυνατόν ένα κτίριο

να έχει ανάγκες για επιπλέον θέρμανση και ψύξη 100% σε ετήσια βάση, σε οποιαδήποτε περιοχή και αν βρίσκεται) συνεπώς αν η ανώτατη τιμή της κλίμακας μας ήταν το 43 θα ήταν σαν να γίνεται παραδοχή για την ύπαρξη του κτιρίου που μόλις αναφέρθηκε, κάτι το οποίο θα ήταν εσφαλμένο.

Αντίστοιχα δεν υπάρχουν τόσο αποδοτικά κτίρια ώστε να ακουμπούν το κάτω όριο του συνόλου τιμών (δηλαδή να μην έχουν καμία απολύτως ανάγκη για θέρμανση και ψύξη και άρα να βρίσκονται στο 0%), συνεπώς αν ήταν το 7 το κατώτατο όριο θα υπήρχε ακριβώς το ίδιο πρόβλημα.

Επιπλέον, με βάση τη συγκεκριμένη κλίμακα ο συντελεστής του κτιρίου αναφοράς ορίζεται στο 16. Η ιδέα πίσω από αυτή την παραδοχή πέρα από πολλές άλλες παραμέτρους, κρύβεται και στην τελευταία γραμμή του πίνακα 3 με τις ενεργειακές κλάσεις. Δεδομένου ότι το μη αποδοτικότερο ενεργειακά κτίριο έχει συντελεστή μέσης καταναλισκόμενης πρωτογενούς ενέργειας 2,73 φορές, ίσως και περισσότερο από το κτίριο αναφοράς και με δεδομένο ότι ο συντελεστής αυτού του κτιρίου (της κατηγορίας Η δηλαδή) κυμαίνεται κοντά στο 44-47, το 16 είναι μία καλή επιλογή για το κτίριο αναφοράς.

Με τη βοήθεια του συντελεστή του κτιρίου αναφοράς ( δηλαδή το 16 ) και με τη βοήθεια του πίνακα 3 προκύπτουν τα εξής όρια για τις ενεργειακές κλάσεις:

Ενεργειακή κλάση A+:  $0 \leq EP \leq 5,28$

Ενεργειακή κλάση A:  $5,28 < EP \leq 8$

Ενεργειακή κλάση B+:  $8 < EP \leq 12$

Ενεργειακή κλάση B:  $12 < EP \leq 16$

Ενεργειακή κλάση C (Γ):  $16 < EP \leq 22,56$

Ενεργειακή κλάση D (Δ):  $22,56 < EP \leq 29,12$

Ενεργειακή κλάση E (Ε):  $29,12 < EP \leq 36,32$

Ενεργειακή κλάση F (Ζ):  $36,32 < EP \leq 43,68$

Ενεργειακή κλάση G (Η):  $43,68 \leq EP$

Βάσει των παραπάνω ανισοτήτων προκύπτουν οι στήλες E, F , η στήλη E αντιστοιχεί την κάθε κλάση σε έναν αριθμό ( από 0 έως 9) . Η στήλη αυτή χρησιμοποιείται παρακάτω στα dataset που θα δεχθεί ο αλγόριθμος ELM σαν είσοδο και για την εκπαίδευση και για τον έλεγχο. Αντίστοιχα, προέκυψαν οι στήλες για τις άλλες δύο πόλεις (Θεσσαλονίκη και Καστοριά) με μόνη διαφοροποίηση τον αριθμό βαθμομερών θέρμανσης και ψύξης.



## 8.1 Πειράματα με τον αλγόριθμο ELM

Πραγματοποιήθηκαν διαφορετικά πειράματα για την κάθε πόλη με τον αλγόριθμο ELM (στο eclipse). Ο πηγαίος κώδικας για τον αλγόριθμο βρίσκεται στο παρακάτω σύνδεσμο: [http://www.ntu.edu.sg/home/egbhuang/elm\\_codes.html](http://www.ntu.edu.sg/home/egbhuang/elm_codes.html) (Java Version). Ο αλγόριθμός δέχεται σαν είσοδο δύο αρχεία .txt το ένα για την εκπαίδευση και το δεύτερο για τον έλεγχο. Συνεπώς για την κάθε πόλη έχουμε δύο αρχεία, που το καθένα βρίσκεται στον αντίστοιχο φάκελο με την ονομασία της πόλης. Αξίζει να σημειωθεί ότι το αρχικό dataset με τις 768 γραμμές χωρίστηκε σε δύο μέρη, όπου το πρώτο θα αποτελούσε το αρχείο εκπαίδευσης και το δεύτερο το αρχείο ελέγχου. Γίνανε πολλές δοκιμές για το πώς ακριβώς θα χωριστούν σε δύο μέρη οι 768 γραμμές, ούτως ώστε να πάρουμε το καλύτερο δυνατό αποτέλεσμα. Καταλήξαμε στο ότι το καλύτερο αποτέλεσμα το παίρνουμε όταν το αρχείο εκπαίδευσης είναι περίπου στο 1/3 γραμμών του συνολικού dataset (δηλαδή γύρω στα 250 δείγματα). Το ίδιο παρατηρήθηκε στα dataset και των τριών πόλεων.

Η πρώτη γραμμή των αρχείων περιέχει τρία νούμερα, το πρώτο δείχνει τον αριθμό των γραμμών, το δεύτερο τον αριθμό των στηλών (οι στήλες των μεταβλητών X1-X8 και η στήλη των ενεργειακών κλάσεων) και το τρίτο τον αριθμό των κλάσεων (χρησιμοποιείται για παράδειγμα για την Ν. Φιλαδέλφεια (Αθήνα) η στήλη E του αρχείου dedomena.xlsx). Τα αποτελέσματα που δίνει ο αλγόριθμος είναι ο χρόνος εκπαίδευσης, η ακρίβεια εκπαίδευσης, ο χρόνος ελέγχου και η ακρίβεια ελέγχου. Τα αποτελέσματα από όλα τα πειράματα συνοψίζονται στον σχήμα 5 στο τέλος της εργασίας.

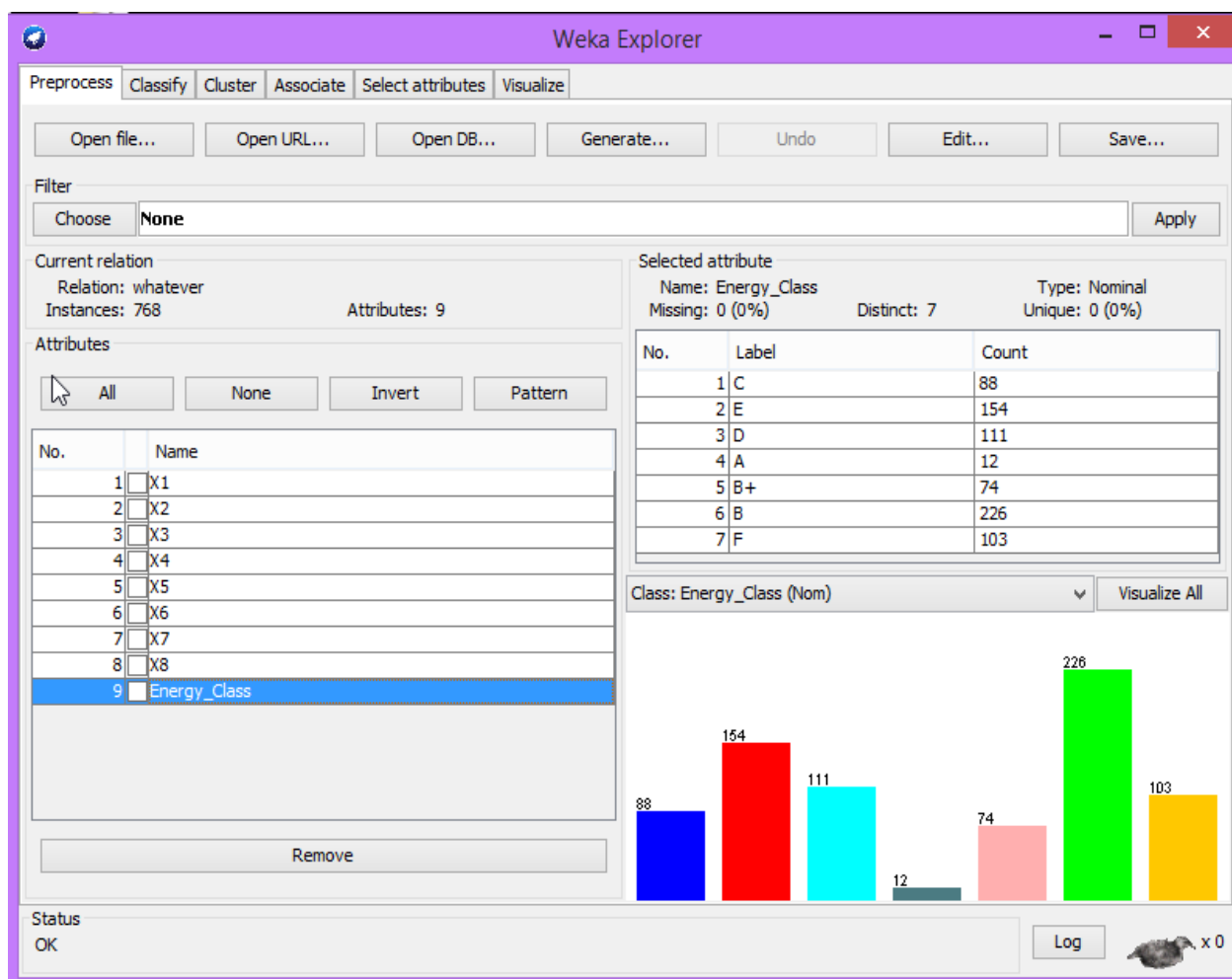
Τα αποτελέσματα του αλγορίθμου για τις τρεις διαφορετικές πόλεις δεν έχουν μεγάλη απόκλιση, ούτε όσον αφορά στους χρόνους εκπαίδευσης, ούτε στην ακρίβεια. Οι χρόνοι εκπαίδευσης κυμαίνονται από 0.094sec – 0.119sec και τα ποσοστά ακρίβειας (δηλαδή σωστής ταξινόμησης) από 84%-87.9%.

## 8.2 Πειράματα με άλλους αλγορίθμους

Για να είναι εφικτό να συγκριθούν τα αποτελέσματα του αλγορίθμου ELM, είναι απαραίτητο να γίνουν πειράματα και με άλλους αλγορίθμους με τα ίδια dataset για να μπορέσουμε να εντοπίσουμε διαφορές στους χρόνους εκπαίδευσης και στα ποσοστά ακρίβειας.

Για το σκοπό αυτό χρησιμοποιήθηκε το weka 3.6, το οποίο δέχεται ως αρχεία εισόδου, αρχεία τύπου .arff, συνεπώς είναι απαραίτητη η μετατροπή των txt dataset σε csv και στη συνέχεια μετατροπή του csv σε arff (Το αρχείο εισαγωγής θα μπορούσε να είναι και csv, αλλά επιλέχθηκε η arff μορφή). Για τη μετατροπή των αρχείων csv σε arff χρησιμοποιήθηκε ο online μετατροπέας που βρίσκεται στον παρακάτω σύνδεσμο: <http://slavnik.fe.uni-lj.si/markot/csv2arff/csv2arff.php>. Το dataset κάθε πόλης για τα συγκεκριμένα πειράματα δεν χωρίστηκε σε δύο μέρη, όπως παραπάνω, συνεπώς έχουμε και τα 768 δείγματα σε ένα αρχείο (συνολικά τρία, ένα για κάθε πόλη)

Στη συνέχεια, και αφού έχουμε πλέον τα αρχεία arff και των τριών πόλεων πραγματοποιείται εισαγωγή του αρχείου arff στο weka. Με την εισαγωγή του αρχείου, στην πρώτη καρτέλα μπορούμε να δούμε τη λίστα με τις 8 μεταβλητές X1-X8 και την ενεργειακή κλάση, τις ελάχιστες, τις μέγιστες και τις μέσες τιμές για κάθε μεταβλητή ξεχωριστά, καθώς και ένα ιστόγραμμα όπου φαίνονται οι συχνότητες των 8 μεταβλητών μαζί συναρτήσει της ένατης. Όλα τα παραπάνω φαίνονται αναλυτικά στην παρακάτω εικόνα:



Εικόνα 4: Εισαγωγή του dataset στο weka

Στη συνέχεια, διαλέγουμε την καρτέλα classify, όπου υπάρχουν τέσσερις διαφορετικές επιλογές, για την πραγματοποίηση των πειραμάτων. Η πρώτη, είναι να χρησιμοποιηθεί ολόκληρο το αρχείο σαν αρχείο εκπαίδευσης, η δεύτερη να οριστεί αντίστοιχα και ένα αρχείο ελέγχου, η τρίτη να ακολουθηθεί η μέθοδος cross validation και η τέταρτη να χωριστεί το εισαγόμενο αρχείο σε δύο μέρη με βάση συγκεκριμένο ποσοστό που επιλέγει ο χρήστης.

Για τα συγκεκριμένα πειράματα, ακολουθήθηκε η μέθοδος του cross validation (δηλαδή μία διαδικασία με την οποία μία μέθοδος που λειτουργεί για ένα δείγμα ενός πληθυσμού

ελέγχεται για εγκυρότητα με την εφαρμογή της μεθόδου σε ένα άλλο δείγμα από τον ίδιο πληθυσμό).

Οι αλγόριθμοι που επιλέχθηκαν για να πραγματοποιηθούν τα πειράματα είναι:

- Από το φάκελο bayes, ο αλγόριθμός Bayes Net και ο αλγόριθμος Bayes Naive
- Από το φάκελο function ο αλγόριθμος LSVM, ο αλγόριθμος RBF Network και ο αλγόριθμος Multilayer Perceptron
- Από το φάκελο meta, ο αλγόριθμος Multiclass Classifier

Για τον καθένα από τους παραπάνω αλγορίθμους με την πραγματοποίησή τους πειράματος αντλούμε αρκετές πληροφορίες, όπως για παράδειγμα τα δείγματα που ταξινομήθηκαν σωστά, τα δείγματα που ταξινομήθηκαν λάθος, το μέσο απόλυτο σφάλμα, τη ρίζα του μέσου τετραγωνικού σφάλματος, το σχετικό απόλυτο σφάλμα, τη ρίζα του σχετικού τετραγωνικού σφάλματος κ.α.

## 9. Αποτελέσματα – Συμπεράσματα

Στο παρακάτω σχήμα συνοψίζονται τα αποτελέσματα από την εκτέλεση των πειραμάτων (αρχείο results.xlsx) για κάθε πόλη και για κάθε αλγόριθμό φαίνεται η ακρίβεια και ο χρόνος εκπαίδευσης της εκάστοτε ταξινόμησης:

A	B	C	D	E	F	G
	ΑΘΗΝΑ		ΘΕΣΣΑΛΟΝΙΚΗ		ΚΑΣΤΟΡΙΑ	
	CORRECTLY CLASSIFIED	TIME	CORRECTLY CLASSIFIED	TIME	CORRECTLY CLASSIFIED	TIME
ELM	84%	0,119	85,6%	0,117	87,9%	0,094
LSVM	70,1%	0,56	64%	0,36	62%	0,42
BAYES NET	78,7%	0,02	79,55%	0,05	78,9%	0,02
BAYES NAÏVE	71,61%	0,01	75,3%	0,01	74,9%	0,01
MULTICLASS CLASSIFIER	65,5%	0,12	70,2%	0,13	72,52%	0,12
RBF NETWORK	88%	2,21	84,8%	2,59	84,8%	3,4
MULTILAYER PERCEPTRON	90,5%	1,17	90,1%	0,95	88,5	0,96

Σχήμα 5: Σύνοψη αποτελεσμάτων από την εκτέλεση των πειραμάτων

Διαπιστώνεται από τα αποτελέσματα, ότι και για τις τρεις πόλεις όλοι οι αλγόριθμοι δίνουν παρόμοια αποτελέσματα με πολύ μικρή απόκλιση.

Τη λιγότερο αποδοτική ταξινόμηση την κάνει ο αλγόριθμος Multiclass Classifier και ο LSVM παρόλο που οι χρόνοι εκπαίδευσης του τελευταίου είναι πολύ καλοί. Την αποδοτικότερη ταξινόμηση την κάνουν οι αλγόριθμοι ELM και RBF Network και Multilayer Perceptron με τη μόνη διαφορά ότι ο αλγόριθμος RBF Network είναι σχεδόν 20 φορές πιο αργός από τον αλγόριθμο ELM και ο αλγόριθμος Multilayer Perceptron δέκα φορές πιο αργός. Αξίζει να σημειωθεί ότι ο αλγόριθμος Multilayer Perceptron, θεωρείται πρόγονος του αλγορίθμου ELM και μάλιστα ο αλγόριθμος ELM προσπαθεί να βελτιώσει τις ατέλειές του και κυρίως τον μεγάλο χρόνο εκπαίδευσης του, πράγμα που επιβεβαιώνεται και από τα νούμερα των πειραμάτων.

Συμπεραίνουμε πως σε ένα συνδυασμό ακρίβειας και ταχύτητας, ο αλγόριθμος ELM υπερέχει έναντι όλων των υπολοίπων, επιτυγχάνει καλύτερη απόδοση γενίκευσης και είναι εξαιρετικά γρήγορος.

Αποδεικνύεται, σε ακόμα ένα πρόβλημα ταξινόμησης, πως ο αλγόριθμος ELM επιτυγχάνει πολύ μεγάλη ακρίβεια ταξινόμησης και μεγάλη ταχύτητα σύγκλισης,

συγκριτικά με τους παραδοσιακούς αλγόριθμους μάθησης ξεπερνώντας αρκετά από τα προβλήματα που αντιμετωπίζουν οι τελευταίοι.

Συνοψίζοντας την θεωρία με την πράξη ο αλγόριθμος ELM φαίνεται να ξεπερνά το πρόβλημα του τοπικού ελάχιστου, τη χρήση αποκλειστικά διαφορίσιμης συνάρτησης ενεργοποίησης, τον ακατάλληλο ρυθμό μάθησης και την υπερπροσαρμογή. Η εξαιρετικά μεγάλη ταχύτητα εκμάθησης του αλγορίθμου ELM, φαίνεται να ξεπερνά το φράγμα του ανώτατου ορίου ταχύτητας πολλών αλγορίθμων εκμάθησης που αποτελεί πρόβλημα ακόμα και σε απλές εφαρμογές με όχι τεράστιο όγκο δεδομένων, ενώ τέλος πολύ σημαντική διαφορά είναι πως οι παραδοσιακοί αλγόριθμοι μάθησης δίνουν βαρύτητα μόνο στην ελαχιστοποίηση του σφάλματος εκπαίδευσης, σε αντίθεση με τον αλγόριθμο ELM, ο οποίος συνυπολογίζει και το ελάχιστο της νόρμας του αθροίσματος των βαρών.

## ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ

Ξενόγλωσσος όρος	Ελληνικός Όρος
feed forward neural network	δίκτυα πρόσθιας τροφοδότησης
Support vector machine	Μηχανή διανυσμάτων υποστήριξης
gradient-based algorithms	αλγόριθμοι που βασίζονται στην κλίση
local minima	τοπικό ελάχιστο
overfitting	υπερπροσαρμογή
Extreme learning machine	Μηχανή μάθησης υψηλής απόδοσης
Hidden nodes	Κρυφοί κόμβοι
Output Matrix	Μήτρα εξόδου
Radial Basis Function networks	Δίκτυα συναρτήσεων βάσης ακτινικού τύπου
Classification problems	Προβλήματα ταξινόμησης
Back Propagation algorithm	Αλγόριθμος πίσω διάδοσης του λάθους
Differential Evolution Algorithm	Αλγόριθμος διαφορικής εξέλιξης
Mutation, crossover, selection	Μετάλλαξη, διασταύρωση, επιλογή
Fitness	καταλληλότητα
classifier	ταξινομητής
Recurrent neural networks	
Complex domain	Πολύπλοκος τομέας
Quadrature amplitude modulation	Διαμόρφωση πλάτους τετραγωνικών σημάτων
Batch learning	Μαζική μάθηση
Chunk by chunk	Κομμάτι- κομμάτι
Fuzzy inference systems	Ασαφή συστήματα συμπερασμού
Regression problems	Προβλήματα παλινδρόμησης
Multiresponse sparse regression algorithm	Αλγόριθμος πολλαπλών αποκρίσεων αραιής παλινδρόμησης
Least angle regression	Αλγόριθμος παλινδρόμησης ελάχιστης γωνίας
Prediction sum of squares	Πρόβλεψη αθροίσματος τετραγώνων
outlier	Ακραία εσφαλμένη τιμή
Empirical risk minimization	Ελαχιστοποίηση εμπειρικού ρίσκου
Standard optimization method	Κλασσική μέθοδος βελτιστοποίησης
Single class classification	Ταξινόμηση απλής κλάσης
Histocompatibility complex molecules	Μόρια συμπλέγματος ιστοσυμβατότητας
Human leukocyte antigens	Αντιγόνα ανθρώπινων λευκοκυττάρων
Zero-crossing rate	Ρυθμός μηδενικής διασταύρωσης
Short-term energy	Βραχυπρόθεσμη ενέργεια
Crest factor	Συντελεστής κορυφής
Root mean square	Μέση τετραγωνική ρίζα
Spectral centroid / spectral flux	Φασματικός κεντροειδής / Φασματική ροή
principal component analysis	ανάλυση σε κύριες συνιστώσες
independent Component analysis	ανάλυση ανεξάρτητων συνιστωσών
Heating load	Φορτίο θέρμανσης
Cooling load	Φορτίο ψύξης
Activation function	Συνάρτηση ενεργοποίησης

## ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ

ELM	Extreme Learning Machine
E-ELM / C-ELM	Evolutionary/ Complex ELM
OS-ELM/ EOS-ELM/ OS Fuzzy ELM	Online Sequential/ Enhanced Online Sequential/ Online Sequential Fuzzy ELM
SCELM	Sine/ Cosine ELM
I-ELM/ CI-ELM/ EI-ELM	Incremental/ Convex Incremental /Enhanced Incremental ELM
OP-ELM/ P-ELM	Optimally Pruned/ Pruned ELM
S-ELM/ SBELM/ R-ELM	Sparse/ Sparse Bayesian/ Regularized ELM
RI-ELM/ V-ELM/ VOP-ELM	Random Iterative/
SLFN/ FNN	Single Layer Feed forward Networks/Feed forward Neural Networks
RBF	Radial Basis Function
SVM - LS-SVM	Support Vector Machine - Least-Squares Support Vector Machine
ISVM/ OSVM	Improved/
ELM_NN/SVM_NN	Extreme Learning Machine/ Support Vector Machine Neural Nets
BP - CBP	Back Propagation - Complex Back Propagation
DE	Differential Evolution
QAM	Quadrature Amplitude Modulation
KNN	K-Nearest Neighbor
CMRAN	Complex Minimal Resource Allocation Network
CRBF	Complex Radial Basis Function
FIS	Fuzzy Inference Systems
FNN	Feedforward neural network
MRSR	Multiresponse Sparse Regression
LOO	Leave One -Out
MLP	Multilayer Perceptron
SCC	Single Class Classification
MC / SCMC / GMC	Mapping Convergence /
HLA	Human Leukocyte Antigens
BCI	Brain Computer Interface

LDA	Linear Discriminant Analysis
ZDA	Z-scale Discriminant Analysis
PCA/ B2DPCA	Principal Component Analysis / Bilateral Two- dimensional PCA
ICA	Independent Component Analysis
OVO	One Versus One
UCI	University of California
K.Ev.A.K	Κανονισμός Ενεργειακής Απόδοσης Κτιρίων
RMS	Root Mean Square
CDH	Cooling Degree Hours
HDD	Heating Degree Days
LED	Light-emitting diode
DNA	Deoxyribonucleic acid
ΘΨΚ	Θέρμανση - Ψύξη - Κλιματισμός
ZNX	Ζεστό νερό χρήσης



## ΑΝΑΦΟΡΕΣ

- [1] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks," *International Joint Conference on Neural Networks (IJCNN'2004)*, vol.2, pp. 985-990, 2004.
- [2] G.-B. Huang, Q.-Y. Zhu and C.-K. Siew, "Extreme Learning Machine: Theory and Applications", *Neurocomputing*, vol. 70, no. 1–3, pp. 489-501, 2006.
- [3] G.-B. Huang and C.-K. Siew, "Extreme learning machine: {RBF} network case," *8th International Conference on Control, Automation, Robotics Vision (ICARCV)*, Kunming, China, vol. 2, pp.1029–1036, 2004
- [4] G.-B. Huang and C.-K. Siew, "Extreme learning machine with randomly assigned RBF kernels," *International Journal of Information Technology & Decision Making*, vol. 11, no. 1, pp. 16–24, 2005.
- [5] G. Huang, L. Chen, and C. Siew, "Universal approximation using incremental constructive feed forward networks with random hidden nodes," *IEEE Transactions on Neural Networks*, vol. 17, no. 4, pp. 879–892, 2006.
- [6] G.-B. Huang, M.-B. Li, L. Chen and C.-K. Siew, "Incremental Extreme Learning Machine With Fully Complex Hidden Nodes," *Neurocomputing*, vol. 71, pp. 576-583, 2008
- [7] G.-B. Huang and L. Chen, "Enhanced Random Search Based Incremental Extreme Learning Machine," *Neurocomputing*, vol. 71, pp. 3460-3468, 2008
- [8] G.-B. Huang and L. Chen, "Convex Incremental Extreme Learning Machine," *Neurocomputing*, vol. 70, pp. 3056-3062, 2007
- [9] M.-B. Li, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "Fully Complex Extreme Learning Machine," *Neurocomputing*, vol. 68, no.1-4, pp. 306-314, 2005.
- [10] Z. Bai, G.-B. Huang, D. Wang, H. Wang and M. B. Westover, "Sparse Extreme Learning Machine for Classification", *IEEE Transactions on Cybernetics*, vol.25, no.4, pp. 836-843, 2014.
- [11] Y. Miche, A. Sorjamaa, and A. Lenda, "OP-ELM: Theory, Experiments and a Toolbox," *Artificial Neural Networks – ICANN*, vol 5163, pp145-154, 2008
- [12] Y. Miche, A. Sorjamaa, P. Bas, O. Simula, C. Jutten, and A. Lendasse, "OP-ELM: optimally pruned extreme learning machine," *IEEE Transactions on Neural Networks*, vol. 21, no. 1, pp. 158–162, 2010.
- [13] J.Luo, C.-M. Vong, P.-K. Wong, " Sparse Bayesian Extreme Learning Machine for Multi-classification", *IEEE Trans. on Neural Network*, vol. 25, no.4 pp. 836 - 843, 2014
- [14] H.-J. Rong, Y.-S. Ong, A.-H. Tan, and Z.Zhu, "A fast pruned-extreme learning machine for classification problem," *Neurocomputing*, vol.72, no.13-15 pp. 359-366, 2008
- [15] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Transactions, Neural Networks*, vol. 17, no. 6, pp. 1411–1423, 2006.
- [16] Y. Lan, Y.-C. Soh, and G.-B. Huang, "Ensemble of online sequential extreme learning machine", *Neurocomputing*, vol 72, pp. 3391-3395, 2009.
- [17] Q.-Y. Zhu, A. K. Qin, P. N. Suganthan, and G.-B. Huang, "Evolutionary extreme learning machine," *Pattern Recognition*, vol. 38, no. 10, pp. 1759–1763, 2005.
- [18] H.-J. Rong, G.-B. Huang, N. Sundararajan, and P. Saratchandran, "Online sequential fuzzy extreme learning machine for function approximation classification problems," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 39, no. 4, pp. 1067–1072, 2009.
- [19] F. Han and D.S. Huang, "Improved extreme learning machine for function approximation by encoding a priori information", *Neurocomputing*, vol.69, no.16-18, pp. 2369-2373, 2006.
- [20] W. Deng, Q. Zheng and L. Chen., "Reguralized Extreme Learning Machine", *Computational Intelligence and Data Mining(CIDM)*, pp. 389 – 395, 2009.
- [21] G.-B. Huang, X. Ding and H. Zhou, "Optimization method based extreme learning machine for classification", *Neurocomputing*, vol.74, no.1-3, pp. 155-163, 2010.
- [22] B.Frenay and M.Verleysen, "Using SVMs with randomized feature spaces: an extreme learning approach", *18<sup>th</sup> European Symposium on Artificial Neural Networks (ESANN)*, pp. 315–320, 2010.
- [23] N.-H. Vo, M.-T. Hoang, H.-T. Huynh, J.-J. Kim and Y. Won, "Iterative extreme learning machine for single class classifier using general mapping convergence framework", *6<sup>th</sup> International Conference on Applied computer science*, pp. 549-554, 2006.
- [24] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.

- [25] R. Zhang, G. Huang, N. Sundararajan, and P. Saratchandran, "Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 3, pp. 485–495, 2007.
- [26] J. Cao and L. Xiong, "Protein Sequence Classification with Improved Extreme Learning Machine Algorithms," *BioMed Research International*, vol. 2014, Article ID 103054, 12 pages, 2014.
- [27] D. Wang and G.-B. Huang, "Protein sequence classification using extreme learning machine," *International Joint Conference on Neural Networks (IJCNN)*, vol. 3, pp. 1406–1411, 2005.
- [28] J. Wang, Z. Yi, J.-M. Zurada, B.-L. Lu and H. Yin, "Extreme learning machine for predicting HLA-Peptide binding, 3<sup>rd</sup> International Symposium on Neural Networks, vol.3973, pp. 716-721, 2006.
- [29] N. Liang, P. Saratchandran, G.-B. Huang and N. Sundararajan, "Classification of mental tasks from EEG signals using extreme learning machine", *International Journal of Neural Systems*, vol.16, no.1, pp. 29-38, 2006.
- [30] J.Siva Prakash and R. Rajesh, "Random iterative extreme learning machine for classification of electronic nose data", *International journal of wisdom based computing*, vol.1, no.3 ,pp. 24-27, 2011.
- [31] R. Minhas, A.-A. Mohammed and Q.-M.-J. Wu, "A fast recognition framework based on extreme learning machine using hybrid object information", *Neurocomputing*, vol. 73, no. 10-12, pp. 1831-1839, 2009.
- [32] R. Minhas, A.Baradarani and S. Seifzadeh, "Human action recognition using extreme learning machine based on visual vocabularies, *Neurocomputing*, vol.73, no.10-12, pp. 1906-1917, 2010.
- [33] R. Minhas, A.-A. Mohammed, Q.-M.-J. Wu and M.-A.-S. Ahmed, "Human face recognition based on multidimensional PCA and extreme learning machine, *Pattern Recognition*, vol. 44, no. 10-11, pp. 2588-2597, 2011.
- [34] A. Tsanas and A. Xifara, "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools", *Energy and Buildings*, vol. 49, pp. 560-567, 2012.
- [35] Z. Yu, F. Haghghat, B.-C.-M. Fung and H. Yoshino, "A decision tree method for building energy demand modeling", *Energy and buildings*, vol.42, pp. 1637-1646, 2010.
- [36] A.-E. Ben-Nakhi and M.-A. Mahmoud, "Cooling load prediction for buildings using general regression neural networks", *Energy Conversion and Management*, vol. 45, pp. 2127-2141, 2004.
- [37] Υπουργείο Περιβάλλοντος Ενέργειας και Κλιματικής αλλαγής -Υ.Π.Ε.Κ.Α. -, Τεχνική Οδηγία Τεχνικού Επιμελητηρίου Ελλάδος, Τ.Ο.Τ.Ε.Ε 20701-1/2010, Αναλυτικές Εθνικές Προδιαγραφές Παραμέτρων Για Τον Υπολογισμό Της Ενεργειακής Απόδοσης Κτιρίων Και Την Έκδοση Του Πιστοποιητικού Ενεργειακής Απόδοσης, 2010
- [38] Υπουργείο Περιβάλλοντος Ενέργειας και Κλιματικής αλλαγής -Υ.Π.Ε.Κ.Α. -, Τεχνική Οδηγία Τεχνικού Επιμελητηρίου Ελλάδος, Τ.Ο.Τ.Ε.Ε 20701-2/2010, Θερμοφυσικές Ιδιότητες Δομικών Υλικών και Έλεγχος της Θερμομονωτικής Επάρκειας των Κτιρίων, 2010.
- [39] Υπουργείο Περιβάλλοντος Ενέργειας και Κλιματικής αλλαγής -Υ.Π.Ε.Κ.Α. -, Τεχνική Οδηγία Τεχνικού Επιμελητηρίου Ελλάδος, Τ.Ο.Τ.Ε.Ε 20701-3/2010, Κλιματικά Δεδομένα Ελληνικών Περιοχών, 2010.
- [40] Υπουργείο Περιβάλλοντος Ενέργειας και Κλιματικής αλλαγής -Υ.Π.Ε.Κ.Α. -, Τεχνική Οδηγία Τεχνικού Επιμελητηρίου Ελλάδος, Τ.Ο.Τ.Ε.Ε 20701-4/2010, Οδηγίες και Έντυπα Ενεργειακών Επιθεωρήσεων Κτιρίων, Λεβήτων και Εγκαταστάσεων Θέρμανσης και Εγκαταστάσεων Κλιματισμού, 2010.
- [41] Dong Li, "Extreme Learning Machines : Basic ELM Algorithms" (JavaVersion), [http://www.ntu.edu.sg/home/egbhuang/elm\\_codes.html](http://www.ntu.edu.sg/home/egbhuang/elm_codes.html)
- [42] UCI, Machine Learning Repository, Center for Machine Learning and Intelligent Systems, Energy Efficient Data Set, <http://archive.ics.uci.edu/ml/datasets/Energy+efficiency>