Third National Conference on Computational Intelligence (NCCI 2019) IOP Publishing
Journal of Physics: Conference Series **1427** (2020) 012012 doi:10.1088/1742-6596/1427/1/012012

Machine learning algorithms for prediction of dyslexia using eye movement

Ms. Vani Chakraborty[1], **Dr. Meentachi Sundaram**[2]

[1] Research Scholar, Department of Computational Sciences and IT, Garden City University, Bengaluru

[2] Assistant Professor, Department of Computational Sciences and IT, Garden City University, Bengaluru

Abstract. Dyslexia is the most widely recognized neurological learning disability. It causes trouble in perusing, composing and spelling. All these can influence scholastic achievement, confidence, and social-emotional development. As roughly 10% of the individuals worldwide are dyslexic, it is a worry of numerous youngsters and grown-ups far and wide. Figuring out how to help the lives of a dyslexic would be of incredible advantage to entire social orders. Studies have demonstrated that the earlier dyslexia is recognized and backing is given in education and training, the more its negative impacts can be alleviated. Subsequently, building up a solid and target screening technique to analyze dyslexia at an early age would be of most extreme significance.

Keywords: Neurological Learning Disability, Dyslexia, Social-emotional development

1. INTRODUCTION

Utilizing an eye-tracker, it is conceivable to record the developments of eyes during different exercises. Following them during perusing is particularly productive on account of dyslexics, as it has been demonstrated that readers with dyslexia have diverse eye developments than typical readers (Rayner 1998). Dyslexics show more and longer obsessions, shorter saccades, and by and large progressively sporadic eye development (Deans et al. 2010; De Luca et al. 2002; Lefton et al. 1979). Learning of this marvel fills in as an important beginning stage in structure a device to isolate ordinary perusers from dyslexics. For this reason, AI gives techniques in recognizing examples and making forecasts dependent on them. Joining the known contrasts among dyslexic and typical eye developments with the element based forecasts given by AI techniques appears to be a characteristic blend to be tried. Applying AI in the discovery of dyslexia from eye developments is a moderately new approach. Rello and Ballesteros (2015), Lustig (2016) and Benfatto et al. (2016) have examined this strategy and gotten promising outcomes. All examinations applied the Support Vector Machine classifier for isolating dyslexics from typical perusers. Lustig also utilized Feed-Forward Neural Networks in the grouping with great outcomes. The majority of the examinations reason that foreseeing the perusing capacity of people from eye developments recorded with eye-following can be effective and dependable. In this postulation, we applied AI in identifying understudies with dyslexia from an enormous 1 eye development accounts information. The information had been gathered for an examination in regards to Internet perusing abilities among understudies with and without learning handicaps. Our objective was to make a product that could dependably foresee people with dyslexia from the given information. Also, we needed to increase understanding on the best way to best utilize the accessible information to recognize dyslexia.

2. REVIEW OF LITERATURE

[1] This paper presents the results of simulations of a brand new category of artificial neural network models of reading. not like previous models, they're not restricted to mono-syllabic words, need no difficult input-output representations like Wickelfeatures and, though supported the NETtalk system of Sejnowski and Rosenberg (1987), need no pre-processing to align the letters and phonemes within the coaching information. the most effective cases square measure able to succeed 100 percent performance on the Seidenberg and McClelland (1989) coaching corpus, in more than ninetieth on rolling nonwords and on harm exhibit symptoms just like nonheritable surface learning disorder.
[2] This paper aims to gift a procedure tool supported neural networks to notice folks with potential learning disabilities (dyslexia). Participants were aged from 9 to eighteen years previous with or while not the diagnosing of learning disorder. Neural networks technique showed consistency in addressing the issues of pattern recognition within the screening of learning disorder.
[3] During this paper a brand new tool is planned as a potential aid to review variations and similarities between the human and also the artificial neural network (NN) learning of some verbal and mathematical elementary skills. For this purpose, easy NNs of the multi-layer kind (MLNN) are build. These MLNNs square measure able to acknowledge some graphemes and/or to form additions of integers up to one thousand. Associate in nursing algorithmic rule supported dynamic character recognition has allowed to limit considerably the info size, creating easier the NN optimisation part of coaching. The adopted methodology of printed symbol coding has allowed to get mechanically massive coaching sets upon that the MLNNs are trained. Then, a check set has been generated to judge the MLNN prediction capability. The analysis of results has shown some fascinating characteristics of the trained nets, such as, as an example, the potential look of terribly rudimentary symptoms analogous to learning disorder. The specialization of the operate of some teams of neurons within the neural system has been conjointly investigated by procuring a synthetic harm to the MLNN (in one or a lot of neurons) and by evaluating the MLNN response.
[4] The lobe within the human brain is liable for low-level audio perception, whereas the pre- frontal lobe takes active role in decoding the audio info. This paper introduces a completely unique approach to grasp the interrelatedness between the temporal and also the pre-frontal lobes of the brain in decoding vowel sounds. The inter-relation is observed by 2 approaches. The primary approach computes correlation live between the direct brain signals of the aforementioned 2 lobes. The upper the correlation, the higher is that the interrelatedness between the activated lobes. The second approach aims at developing a feature-level mapping between the temporal and also the lobe brain activations. The motivation of the second approach lies in examining the uniformity within the learnt neural weights once convergence for a similar vowel audio stimulant no matter the diurnal variations within the brain signals.

Though any ancient mapping functions can be utilised to undertake the temporal to anterior mapping, we tend to used a type-2 fuzzy neural network to serve the aim. Experiments undertaken ensure that the weights of the planned type-2 fuzzy neural internet converges quicker than its type-1 counterpart and back-propagation neural network. The quicker convergence of weights represent that the planned type-2 fuzzy neural network captures higher audio sensory activity ability than the remainder. The planned work is predicted to seek out applications within the early detection of disorder in auditive perceptual-ability, sometimes said as learning disorder.

3. DATA COLLECTION

The information utilized for this examination was given by the e-Seek venture bunch from the

Department of Psychology at the University of Jyväskylä. Their examination task was about Internet perusing aptitudes among Finnish understudies with and without learning incapacities. The information had been gotten through the span of three years and contains information of 165 youths with a normal period of 12.5 years: their aftereffects of the tests done, eye-development information, and incomplete examination of these. The understudies had been looked over a class of around 400 understudies. Of the picked understudies, 30(18%) met the criteria for a perusing issue dependent on picking the tenth most noticeably awful percentile of the perusing fluency execution score. This criteria was utilized to mark the understudies as either dyslexic or ordinary perusers. At the point when contrasted with the general commonness of dyslexia built up in segment 2.1, the dyslexics in this information are marginally over-spoke to.

The eye developments of the members were recorded utilizing an Eye Link 1000 (SRResearch, manual) eye-tracker with an examining recurrence of 1000 Hz. A Dell Precision T5500 workstation with an Asus VG-236 screen (1920 x 1080, 120 Hz, 52 x 29 cm) at the survey separation of 60cm was utilized for showing the boosts. The alignment of the gadget was performed before the experiment and rehashed between preliminaries, if unmistakable head developments were made, a float was identified on the scientist's screen utilized for following the eye developments, or the adjustment mistake surpassed .30 visual degrees. (Hautala et al. 2018).

During the experiment, members finished a training undertaking and afterward 10 recreated data search assignments. The assignments comprised of perusing the contextualized question and after that choosing a query item (out of four alternatives), which would enable them to respond to the inquiry. A case of the given inquiry is "Discover why pandas are imperiled?" (Hautala et al. 2018)

4.DATA PREPROCESSING

The got information is organized by containing one fixation for every column. The 10 assignments and the training task that the members finished are consequently alluded to as preliminaries for the good of clarity, as this is their name in the information file.

The eye development information was likewise tidied up before being utilized for making the capabilities in this examination. The SPSS Statistics (IBM Corporation) programming was utilized for the preprocessing. The means led are beneath:

- The members with no perusing fluency execution score were forgotten about, in light of the fact that this score was required in building up whether the member had an understanding issue or not.
- The information of the training preliminary was expelled, as it was centered around preparing the members for the real errands.
- The first fixation of every preliminary was evacuated because of the following being off base at this stage. • The lines that contained the worth "1" in the Bad Data segment were expelled. This mark showed that the specific information line contained awful information.
- Participant with the id 396 was evacuated because of missing an excess of pertinent information after the above activities.

After the preprocessing, 161 understudies were left in the information file. Of these, 30 are perceived as having a perusing issue dependent on their perusing fluency execution score.

5.MACHINE LEARNING METHODS

The picked AI techniques for this paper is Support Vector Machine and Random Forest. These were chosen since they are at present broadly utilized strategies. Bolster Vector Machine has likewise been recently utilized in investigations of this field (Benfatto et al. 2016; Relloand Ballesteros 2015;Lustig 2016),so this was picked to build up the gauge results.

5.1 SUPPORT VECTOR MACHINE

Bolster Vector Machine (SVM), displayed by Cortes and Vapnik (1995), is a broadly utilized and powerful classification strategy. It has effectively been applied in face identification (Osuna, Freund, and Girosit1997) and content acknowledgment (Wang, Babenko, and Belongie 2011), among different issues (Noble et al. 2004; Mountrakis, Im, and Ogole 2011; Noble et al. 2004). SVM isolates classes by mapping the info vectors into a high dimensional element space through the picked non-direct mapping (Cortes and Vapnik 1995). In this space an ideal hyper plane is found for the divisible classes. Figure 1 demonstrates the ideal hyper plane and its edges. This hyperplane is defined as the straight choice capacity with a maximal edge between the information purposes of the two classes. Amplifying this edge has been demonstrated to decrease the speculation blunder (Vapnik 1999).

Assume we are given a lot of preparing information D, with n information focuses

$$D = \{(x_i, y_i) | x_i \in \mathbf{R}^p, y_i \in \{-1, 1\}\}_{i=1}^{n} \qquad (4.1)$$

where x_i is a p-dimensional real number and y_i the class which x_i belongs to (either 1 or -1).

The data points are said to be linearly separable if

$$y_i(w \cdot x_i + b) \geq 1, \ \forall i = 1, \ldots, n \qquad (4.2)$$

where w is a vector perpendicular to the hyperplane and $\frac{b}{||w||}$ the offset of the hyperplane

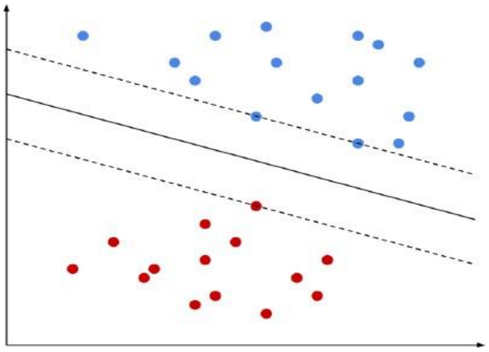


Figure 1: SVM with maximum and margin hyperplane

hyperplane isolating two classes. The ran lines manner that there are information focuses between them. The separation between the edges, which is 2 kwk , would then be able to be attempted to be boosted. Amplifying this separation includes limiting kwk. Subsequently, this is the streamlining issue, with the imperative.

5.2 RANDOM FORESTS

Random Forests (RF) have been utilized with great outcomes in different errands, including information mining (Verikas, Gelzinis, and Bacauskiene 2011), bioinformatics and computational science (Boulesteix et al. 2012), and remote detecting (Belgiu and Dr˘agu¸t 2016). Arbitrary Forest is a classifier that comprises of a gathering of randomized choice trees, which decision in favor of the most prevalent class (Breiman 2001).

The choice tree classifier comprises of an established tree, which contains hubs t0; :::; tn;n 2 N that each speak to a subspace Xtn X. The root hub t0 relates to the info space X. Every hub t is marked with a split st . The parts partition the hubs' subspace Xt into two subspaces, which are spoken to by the hubs' youngsters. (Louppe 2014)

Officially, a Random Forest is characterized (Breiman 2001) as a classifier comprising of an accumulation of tree organized classifiers fhk(x;Tk); k = 1; : :g, where Tk are free indistinguishably circulated arbitrary vectors, and each tree makes a unit choice for the most prominent class at info x. The decency of the choice tree classifier parts is determined by the polluting influence measure, likewise called the Gini file (Alpaydin 2014). As per Alpaydin, a split in the tree is unadulterated "if after the split, for all branches, every one of the occasions picking a branch have a place with a similar class". For more subtleties on the Gini list, see Louppe (2014). A piece of the client customizable parameters for Random Forest include controlling when the parting of the hubs is halted (Louppe 2014). This is essential to anticipate overfitting of the model. The accompanying parameters are utilized to control when a hub t is set as a terminal Figure 2. A choice tree made for a parallel order issue. The tree contains five hubs, of which three (t2; t3; t4) are terminal hubs. Two parts segment the information space into three subspaces. (Figure propelled by Louppe (2014))

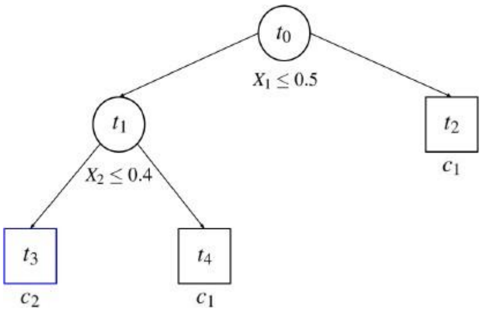


Figure 2: A decision tree created for binary classification problem

Furthermore, when choice trees are incorporated with an irregular backwoods, two additional parameters become pertinent. The quantity of trees in the timberland is characterized by the n_estimators parameter. Having a bigger number of trees is normally better, however that likewise builds the calculation time for the model. When parting a hub in the choice tree, the element utilized for the split is chosen from an arbitrary subset of highlights. The measure of highlights picked into this subsetis dictated by the max_features parameter. (Louppe 2014)

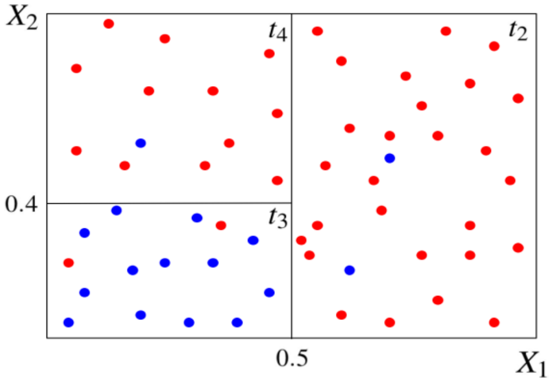


Figure 3: The partition caused in splits by decision tree

Figure 3. The allotments brought about by the parts in choice tree 5. Red dabs speak to objects of class c1 while blue spots speak to objects of class c2. (Figure enlivened by Louppe (2014)) Decision trees and other tree-based strategies are rewarding as clarified by Louppe (2014), on the grounds that they:

- are non-parametric,
- intrinsically execute highlight choice,
- are hearty to exceptions or mistakes in marks,
- Handle heterogeneous information (requested or absolute factors, or a blend of both).

Also, Cutler, Cutler, and Stevens (2012) express that Random Forests are engaging, on the grounds that they:

- Naturally are fit for relapse and order,
- have an implicit gauge of speculation blunder,
- can be utilized straightforwardly for high-dimensional issues.

6.THE IMPLEMENTATION

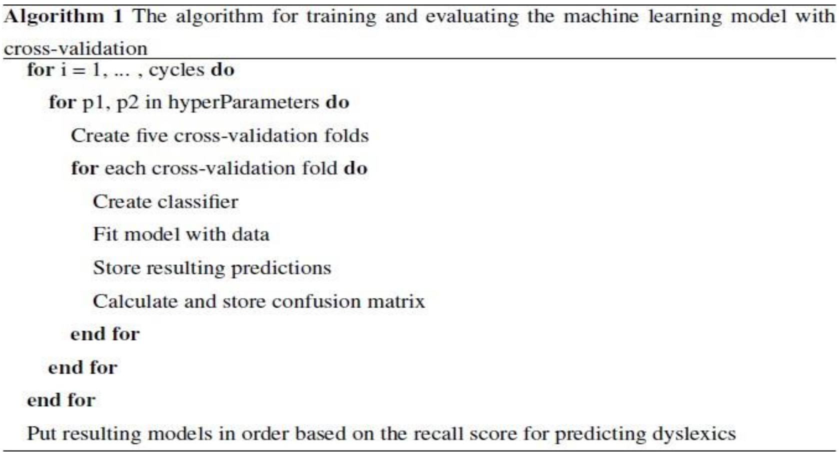
The content made for delivering our outcomes has the accompanying stages: instatement of the information, highlight extraction and age, preparing and assessment of models, and showing results. The instatement of the information contains a couple of significant assignments. The needed sections from the entire information are chosen utilizing Pandas. At the same time, the cells containing obscure qualities are changed into cells containing a whitespace. These cells are then supplanted to cells containing "0". The network used to store the removed information is additionally arranged relying upon the list of capabilities to be made.

The piece of the content utilized for preparing the model and getting the expectation results utilizes a self-made cross-approval and lattice search strategy. This technique is appeared in Algorithm 1. The preparation and assessment are done in cycles; each cycle comprises of preparing the model and getting the outcomes. The cycles steady decides how frequently the entire cross approval cycle is finished. p1 and p2 are two hyperparameters picked to be enhanced. For SVM, C and gamma were improved. Individually, for Random Forest, max_features and n_estimators were chosen to be upgraded. For each new hyperparameter mix, the model is made again to guarantee that it doesn't contain any memory of past trainings.

For our investigation, we chose to utilize 5-overlay cross-approval, since more creases would have decreased the measure of understudies with dyslexia in each overlap to a too limited quantity.

Moreover, utilizing fever folds was computationally quicker. The entire cross- approval procedure is rehashed multiple times to lessen the impact of irregularity on the outcomes. A comparable methodology was utilized in the examination by Mantau et al. (2017), where a 5-overlap cross-approval was utilized to conclude the best parameters, and the entire preparing and testing procedure was rehashed multiple times. This was done to assess the presentation of the models.

The mixes of hyperparameters are looked at against one another by a presentation metric. For our situation, we utilized the review score of dyslexics anticipated. Review is the portion of effectively anticipated examples out of the considerable number of tests of the positive class. This was picked as the presentation metric in this examination as it was regarded more critical to accurately distinguish the dyslexics than typical perusers. Notwithstanding the review score, we likewise watched the general exactness of the model. Utilizing just the exactness score isn't sufficient, on the grounds that the classes are lopsided in our information. It is conceivable to get an exactness of 81.5% by simply announcing the majority of the guineas pigs as typical perusers. This would give a bogus image of the model's exhibition.



Algorithm 1 The algorithm for training and evaluating the machine learning model with cross-validation

```
for i = 1, ... , cycles do
    for p1, p2 in hyperParameters do
        Create five cross-validation folds
        for each cross-validation fold do
            Create classifier
            Fit model with data
            Store resulting predictions
            Calculate and store confusion matrix
        end for
    end for
end for
Put resulting models in order based on the recall score for predicting dyslexics
```

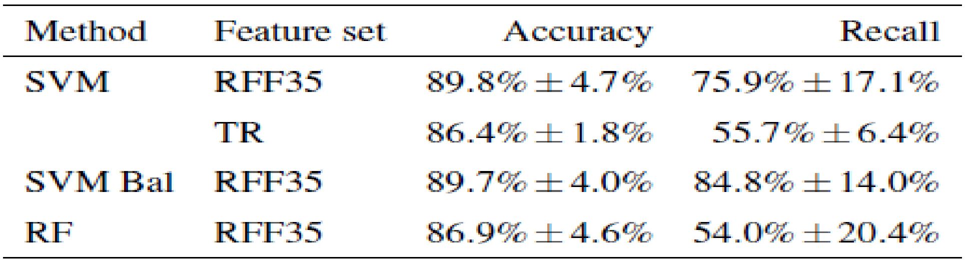
7. RESULTS AND DISCUSSIONS

This section talks about the outcomes got in this examination. These outcomes were altogether accomplished by utilizing the technique depicted before in part 5 with 100 cycles and a 5– overlay cross-approval. The scores exhibited here for each model are midpoints of the 100 cycles. The blunder given is the standard deviation of these outcomes. Table 1 demonstrates a review of the best outcomes. The "Strategy" segment demonstrates the AI technique used to create the model. The "Bal" tag shows that the class loads were adjusted for the Scikit-learn library SVM by modifying them contrarily in extent to class frequencies. The subsequent segment holds the names of the capabilities as given in segment 5.2. The "Exactness" section holds the normal division of right expectations for the majority of the hundred models made during the calculation. The mistake given is the standard deviation of these precision scores. The last segment contains the normal review scores for the class of dyslexics of the hundred models made.

Table 1. Best models made with their exactness and review scores. These are the normal outcomes more than 100 cycles. In this investigation, we set out to create an AI model able to do dependably recognizing understudies with perusing issue utilizing eye development information. The AI calculations chose were the Support Vector Machine and Random Forest. Utilizing the structure science standards in an iterative style, we had the option to accomplish our objective and furthermore acquire information in regards to the issue. The best model, with an exactness of 89.8%, was

accomplished with the Support Vector Machine by utilizing a list of capabilities made from the most significant eye development highlights chosen by Random Forest. This examination indicated promising outcomes in having the option to distinguish understudies with perusing issue dependent on their eye developments. A more straightforward experiment could be set up with the members focusing on perusing more content. From the eye development information got thusly, a comparable arrangement of the most significant highlights could be removed by utilizing the histogram grid list of capabilities and Random Forests include significance estimation. The model made along these lines could then be utilized as a pre-screening instrument for dyslexia location.

Table 1. Best models created with their accuracy and recall scores. These are the average results over 100 cycles.



Method	Feature set	Accuracy	Recall
SVM	RFF35	89.8% ± 4.7%	75.9% ± 17.1%
	TR	86.4% ± 1.8%	55.7% ± 6.4%
SVM Bal	RFF35	89.7% ± 4.0%	84.8% ± 14.0%
RF	RFF35	86.9% ± 4.6%	54.0% ± 20.4%

8.CONCLUSION

In this examination, we set out to deliver an AI model able to do dependably distinguishing understudies with perusing issue utilizing eye development information. The AI calculations chose were the Support Vector Machine and Random Forest. Utilizing the structure science standards in an iterative design, we had the option to accomplish our objective and furthermore acquire information in regards to the issue. The best model, with a precision of 89.8%, was accomplished with the Support Vector Machine by utilizing a list of capabilities made from the most important eye development highlights chosen by Random Forest.

REFERENCES

[1] J.A. Bullinaria, "Neural network models of reading multi-syllabic words", Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan), Vol.1, 1993.

[2] Macário Costa, Jorge Zavaleta, Sérgio Manuel Serra da Cruz, Mary Manhães, Renato Cerceau, Luís Alfredo Carvalho and Renata Mousinho, "A computational approach for screening dyslexia", 26th IEEE International Symposium on Computer-Based Medical Systems, 2013.

[3] Nicola Pio Belfiore, Imre J. Rudas and Apollonia Matrisciano, "Simulation of verbal and mathematical learning by means of simple neural networks", 9th International Conference on Information Technology Based Higher Education and Training (ITHET), 2010.

[4] Mousumi Laha, Amit Konar, Pratyusha Rakshit, Susmita Chaki and Atulya K. Nagar,"Understanding the Biological Underpinning of Auditory Perception for Vowel Sounds Using a Type-2 Fuzzy Neural Network", IEEE Symposium Series on Computational Intelligence (SSCI), 2018.

[5] Belgiu, Mariana, and Lucian Dr˘agu¸t. 2016. “Random forest in remote sensing: A review of applications and future directions”. ISPRS Journal of Photogrammetry and Remote Sensing 114:24–31.

[6] Benfatto, Mattias Nilsson, Gustaf Öqvist Seimyr, Jan Ygge, Tony Pansell, Agneta Rydberg, and Christer Jacobson. 2016. “Screening for dyslexia using eye tracking during reading”. PloS one 11 (12): e0165508.

[7] Bergstra, James, and Yoshua Bengio. 2012. “Random search for hyper-parameter optimization”. Journal of Machine Learning Research 13 (Feb): 281–305.

[8] Boser, Bernhard E, Isabelle M Guyon, and Vladimir N Vapnik. 1992. “A training algorithm for optimal margin classifiers”. In Proceedings of the fifth annual workshop on Computational learning theory, 144–152. ACM.

[9] Boulesteix, Anne-Laure, Silke Janitza, Jochen Kruppa, and Inke R König. 2012. “Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics”. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2 (6): 493–507.

[10] Breiman, Leo. 2001. “Random forests”. Machine learning 45 (1): 5–32.

[11] Chang, Chih-Chung, and Chih-Jen Lin. 2011. “LIBSVM: a library for support vector machines”. ACM transactions on intelligent systems and technology (TIST) 2 (3): 27.

[12] Claesen, Marc, and Bart De Moor. 2015. “Hyperparameter search in machine learning”. arXiv preprint arXiv:1502.02127.