

## Komparasi algoritma *Conjugate gradient* dan *Gradient descent* pada MLPNN untuk Tingkat Pengetahuan Ibu (Studi kasus Pemberian ASI Eksklusif)

Finki Dona Marleny<sup>1)</sup>

<sup>1,2)</sup>STIKES Sari Mulia Banjarmasin

Jl. Pramuka No.02 Telp. (0511) 3268105/Fax (0511) 3270134 Banjarmasin

e-mail: <sup>1)</sup>finkidona@gmail.com,

### Abstrak

Tingkat pengetahuan ibu dalam pemberian ASI eksklusif memiliki peran dalam melaksanakan program pemerintah untuk memberikan ASI Eksklusif selama enam bulan pertama kehidupan. Faktor-faktor berdasarkan tingkat Pendidikan, pekerjaan, umur, jumlah anak, serta dukungan suami dapat menentukan keberhasilan tingkat pemberian ASI secara eksklusif. Dari tingkat pengetahuan ibu dalam pemberian ASI ini dapat diklasifikasikan menggunakan metode klasifikasi Karena data yang berkaitan dengan tingkat pemberian ASI cukup tersedia di lembaga kesehatan seperti Puskesmas di tiap daerah dan kota, namun perlu suatu proses untuk mendapatkan informasi yang dapat digali lagi pengetahuannya agar dapat lebih berguna. Proses ini dapat menggunakan metode klasifikasi. Dalam penelitian ini Multilayer Perceptron Neural network digunakan untuk mengklasifikasi dengan membandingkan optimasi algoritma *Conjugate gradient* dan *Gradient descent*. Hasil yang lebih baik didapatkan dari arsitektur jaringan MLPNN 2 hidden layer, fungsi aktivasi output layer sigmoid, optimasi algoritma *Gradient descent*, hasil yang didapat sebesar 55,9% untuk akurasi prediksi training serta 46,9% untuk akurasi prediksi testing.

**Kata kunci:** eksklusif, ibu, pengetahuan, klasifikasi

### 1. Pendahuluan

Anjuran pedoman internasional untuk pemberian ASI secara eksklusif mengacu pada resolusi World Health Assembly [1], yang didasarkan pada bukti ilmiah tentang manfaat ASI bagi daya tahan tubuh bayi, pertumbuhan dan perkembangannya. Banyak manfaat yang didapat dengan pemberian ASI Eksklusif selama 6 bulan pertama. Di Indonesia berlandaskan keputusan Menteri kesehatan RI No.450/Men.Kes/SK/IV/2004, dikatakan, untuk mencapai pertumbuhan, perkembangan dan kesehatan optimal, bayi harus diberi ASI eksklusif selama 6 bulan pertama, selanjutnya untuk kecukupan nutrisi bayi mulai diberi makanan pendamping ASI yang cukup dan aman, dengan pemberian ASI dilanjutkan sampai usia 2 tahun[2].

Faktor keberhasilan dan kegagalan dalam pemberian ASI secara eksklusif dapat di klasifikasi Karena banyak faktor yang dapat mempengaruhi kegagalan dan keberhasilan dalam pemberian ASI secara eksklusif. Kegagalan menyusui juga disebabkan karena faktor status gizi ibu sebelum hamil, selama hamil dan selama menyusui. Hal ini terjadi karena selama menyusui, terjadi mobilisasi lemak tubuh ibu untuk memproduksi ASI dan simpanan lemak ibu dengan status gizi kurus lebih rendah dari simpanan lemak tubuh pada ibu normal[3].

Hasil Riset Kesehatan Dasar tahun 2010 persentase pola menyusui pada bayi umur 0 bulan adalah 39,8% menyusui eksklusif, 5,1% menyusui dominan, dan 55,1% menyusui parsial. Persentase menyusui eksklusif semakin menurun dengan meningkatnya kelompok umur bayi. Pada bayi yang berumur 5 bulan menyusui eksklusif hanya 15,3%, menyusui dominan 1,5% dan menyusui parsial 83,2% [4].

Tingkat pendidikan merupakan salah satu aspek sosial yang umumnya berpengaruh pada tingkat pendapatan keluarga sebagai faktor ekonomi. Pendidikan juga dapat mempengaruhi sikap dan tingkah laku manusia. Semakin tinggi tingkat pendidikan ibu, semakin tinggi jumlah ibu tidak memberikan ASI pada bayinya. Hal ini mungkin disebabkan karena ibu berpendidikan tinggi biasanya mempunyai banyak kesibukan di luar rumah, sehingga cenderung meninggalkan bayinya. Sedangkan ibu berpendidikan rendah lebih banyak tinggal di rumah sehingga lebih banyak mempunyai kesempatan untuk menyusui

bayinya [2]. Tingkat ekonomi keluarga merupakan salah satu faktor yang menentukan pola pemberian ASI. Di daerah pedesaan keadaan ini cukup nyata, makin tinggi tingkat ekonomi makin berkurang prevalensi menyusui. Namun di negara-negara industri frekuensi menyusui lebih tinggi di kalangan tingkat sosial atas [5].

Dalam penelitian sebelumnya, faktor-faktor dari tingkat pengetahuan ibu dalam pemberian ASI akan dimodelkan dalam klasifikasi prediksi, dengan mendeskripsikan peranan faktor-faktor yang mempengaruhi ibu dalam pemberian ASI. Beberapa faktor yang akan menjadi atribut untuk proses klasifikasi seperti umur, tingkat pendidikan dan pekerjaan dapat dibagi menjadi beberapa kategori yang nantinya akan mengikuti pola dari algoritma klasifikasi seperti tingkat pendidikan di bagi menjadi lulusan SD, SMP, SMA dan Sarjana, kemudian pekerjaan dapat di kategorikan sebagai pekerja ibu rumah tangga (IRT), Swasta dan PNS. Berbagai faktor yang mempengaruhi seorang ibu dalam memberikan ASI inilah yang akan di klasifikasi untuk kemudian di prediksi sehingga membentuk kelompok-kelompok yang memiliki kesamaan tersembunyi dari data yang telah dianalisa [6].

Klasifikasi merupakan suatu proses menemukan kumpulan pola atau fungsi yang mendeskripsikan serta memisahkan kelas data yang satu dengan yang lainnya untuk menyatakan objek tersebut masuk pada kategori tertentu yang sudah ditentukan[7]Klasifikasi data yang dapat digunakan untuk mengetahui tingkat hubungan antar data tersebut adalah dengan metode klasifikasi [8]. Metode-metode klasifikasi antara lain C4.5, *RainForest*, *Naïve Bayesian*, *neural network*, *genetic algorithm*, *fuzzy*, *case-based reasoning*, dan *k-Nearest Neighbor* [9], [10].

Tujuan dari penelitian ini adalah untuk mengetahui tingkatan akurasi klasifikasi serta hubungan atribut yang berkaitan dengan pengetahuan ibu dalam pemberian ASI dengan metode klasifikasi, untuk menganalisa data tingkat pengetahuan ibu dengan membandingkan hasil kinerja optimasi algoritma NN yaitu algoritma *Conjugate gradient* dan *Gradient descent*.

## 2. Metode Penelitian

### 2.1 Pengumpulan dan pengolahan data awal

Data yang akan digunakan merupakan data responden dari pasien puskesmas S.Parman, Sei Mesa & Gambut, dari data tersebut terdapat atribut umur, pekerjaan ibu dan pendidikan. Dari atribut tersebut dapat ditetapkan sebagai default, yaitu dengan kategori sebagai berikut:

Tabel 1. Data awal

Umur	Pekerjaan	Pendidikan terakhir	Anak ke-	Dukungan Suami
<=20 thn, <=30 thn <= 40 thn dan <45 thn	Ibu Rumah Tangga (IRT) Swasta PNS	SD SMP SMA, Sarjana	1,2,3,>=4	Ya, Tidak

### 2.2 Model yang diusulkan

Model yang diusulkan menggunakan metode klasifikasi algoritma *Multilayer Perceptron Neural Network* dengan optimasi algoritma *Conjugate gradient* dan *Gradient descent*, yang akan dianalisa dengan SPSS.

#### 2.2.1 Multilayer Perceptron (MLP)

*Perceptron* yang memiliki satu lapisan bobot hanya memiliki fungsi linear perkiraan input dan tidak bisa memecahkan masalah seperti *XOR*, di mana diskriminan yang akan diestimasi adalah nonlinear. Demikian pula, *perceptron* tidak dapat digunakan untuk regresi nonlinear. Keterbatasan ini tidak berlaku untuk jaringan dengan lapisan menengah atau tersembunyi antara lapisan hidden input dan lapisan output. Jika digunakan untuk klasifikasi, *multilayer* seperti *perceptrons* MLP dapat melaksanakan diskriminan nonlinier dan, jika digunakan untuk regresi, dapat menggunakan perkiraan fungsi nonlinier input [11].

#### 2.2.2 Fungsi Aktivasi

Beberapa fungsi aktivasi yang digunakan adalah sebagai berikut [12]:

- a. Fungsi *Sigmoid*

$$f_{sig}(net_i) = \frac{1}{1 + e^{-s_i \cdot net_i}} \quad (1)$$

- b. Fungsi *Identitas*

$$f(x) = x \quad (2)$$

#### 2.2.3 Gradient Descent with Momentum

---

Algoritma standard yang digunakan dalam pelatihan *MLP Neural Network* adalah *gradient descent with momentum*, perubahan bobot didasarkan atas gradien yang terjadi untuk pola yang dimasukkan saat itu. Modifikasi yang dapat dilakukan adalah melakukan perubahan bobot yang didasarkan atas arah gradien pola terakhir dan pola sebelumnya (disebut momentum) yang dimasukkan, jadi tidak hanya pola masukan terakhir saja yang diperhitungkan. Penambahan momentum dimaksudkan untuk menghindari perubahan bobot yang mencolok akibat adanya data yang sangat berbeda dengan yang lain (outlier). Apabila beberapa data terakhir yang diberikan ke jaringan memiliki pola yang serupa (berarti arah gradien sudah benar), maka perubahan bobot dilakukan secara cepat. Namun, apabila data terakhir yang dimasukkan memiliki pola yang berbeda dengan pola sebelumnya, maka perubahan dilakukan secara lambat. Dengan Penambahan momentum, bobot baru pada waktu ke (t+1) didasarkan atas bobot pada waktu t dan (t-1). Disini harus ditambahkan dua variabel baru yang mencatat besarnya momentum untuk dua iterasi terakhir. Jika  $\mu$  adalah konstanta ( $0 \leq \mu \leq 1$ ) yang menyatakan parameter momentum maka bobot baru dihitung berdasarkan persamaan :

$$w_{:,i}(t+1) = w_{:,i}(t) + \alpha \delta_i z_i + \mu (w_{:,i}(t) - w_{:,i}(t-1))$$

dan

$$v_{:,i}(t+1) = v_{:,i}(t) + \alpha \delta_i x_i + \mu (v_{:,i}(t) - v_{:,i}(t-1)) \quad (3)$$

Dengan cara ini, apabila learning rate terlalu tinggi dan mengarah ke ketidak stabilan, maka learning rate akan diturunkan. Sebaliknya jika learning rate terlalu kecil untuk menuju konvergensi, maka learning rate akan dinaikkan. Dengan demikian, maka algoritma pembelajaran akan tetap terjaga pada kondisi stabil.

#### 2.2.4 Conjugate gradient

*Conjugate gradient* seperti *gradient descent* juga menggunakan *gradien* dari fungsi kinerja untuk menentukan pengaturan bobot-bobot dalam rangka meminimumkan fungsi kinerja. Hanya saja *conjugate gradient*, pengaturan bobot tidak selalu dengan arah menurun tetapi disesuaikan dengan arah konjugasinya [11]. *Conjugate gradient* menggunakan pendekatan penemuan bobot optimal sepanjang arah gradien dengan fungsi line search untuk mencari arah gradien fungsi kinerja. Fungsi *line search* tersebut digunakan untuk menempatkan titik minimum sehingga dapat meminimumkan fungsi kinerja *mean square error*(MSE) selama arah pencarian. Untuk mencari arah pencarian ada 3 formula yang salah satunya adalah formula *Fletcher* dan *Reeves* [12]:

$$\beta = \frac{g_{t+1}^T [g_{t+1} - g_t]}{g_1^T g_t} \quad (4)$$

#### 2.2.5 ROC Curve

*Receiver Operating Characteristic (ROC) Curve* adalah alat untuk mengevaluasi kinerja klasifikasi biner. Fungsinya adalah untuk menentukan ambang batas klasifikasi optimal yang memaksimalkan akurasi klasifikasi dan meminimalkan kesalahan klasifikasi. ROC kurva bergantung pada notasi sensitivitas dan spesifisitas. Nilai-nilai ini tergantung pada set data tertentu [13]. ROC kurva dibangun dengan mengukur dua nilai, yaitu dimulai dengan batas minimum dan kemudian meningkat perlahan-lahan. Kekhususan dan sensitivitas pasangan mampu membenarkan threshold optimal untuk melakukan klasifikasi. ROC kurva memberikan gambaran kinerja keseluruhan tes. Ketika kurva ROC dari tes yang berbeda dibandingkan, kurva klasifikasi baik terletak dekat dengan sudut kiri atas, sementara pengklasifikasi buruk cenderung berkumpul di dekat garis, dihitung sebagai berikut:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (5)$$


---

$$\text{Sensitivity} = \text{TN} / (\text{TN} + \text{FP})$$

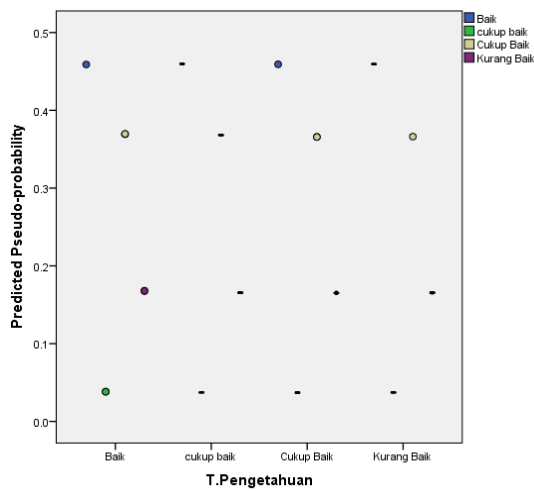
(6)

dimana TP, TN, FP dan FN yang benar-diklasifikasikan kasus positif, kasus negatif benar-diklasifikasikan, kasus positif salah-salah diklasifikasikan dan digolongkan kasus negatif, masing-masing. Secara teori, sensitivitas dan spesifisitas independen dalam arti bahwa adalah mungkin untuk mencapai 100% di kedua. Ambang terbaik dicapai ketika sensitivitas mendekati satu dan spesifisitas mendekati nol. Indikasi lain dari klasifikasi yang baik adalah ketika area di bawah kurva ROC (AUC) dekat 100%.

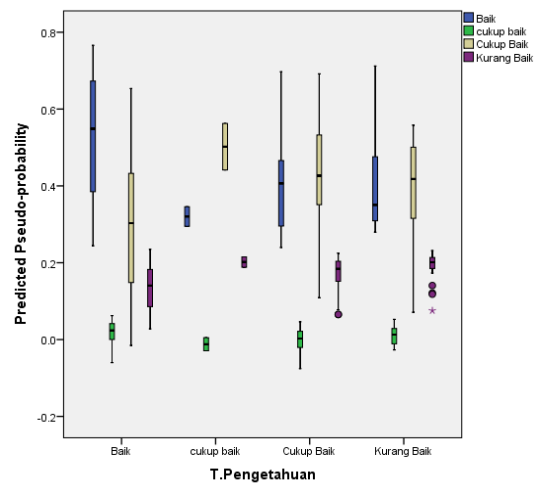
### 2.3. Pengujian Model

Sebanyak 70% data digunakan untuk pelatihan dan 30% untuk uji akurasi. Berikut ini merupakan pengujian model dari arsitektur jaringan MLPNN, yaitu dapat dilihat pada gambar berikut:

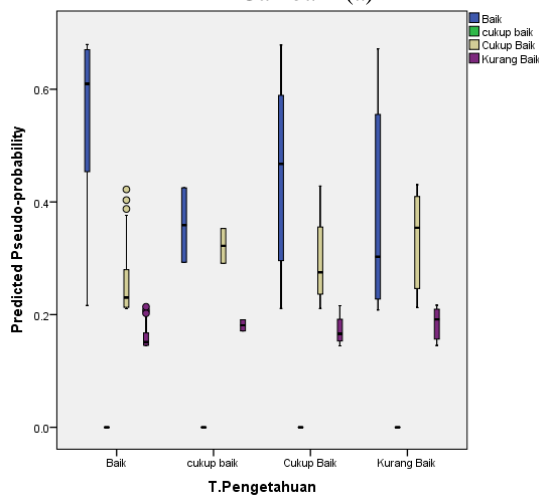
Pada gambar 1(a),(b),(c),(d) merupakan boxplots yang dihasilkan dari arsitektur jaringan neural network dari table 2, untuk membandingkan dataset pada setiap data pada gambar 1, (a),(b),(c) dan (d) adalah boxplot dari tingkat rata-rata, koefisien kluster rata-rata, dan transitivitas, masing-masing hasil dari klasifikasi. *Boxplots* ini merupakan mekanika statistik yang digunakan untuk penetapan fitur vektor untuk klasifikasi melalui MLP Neural Network.



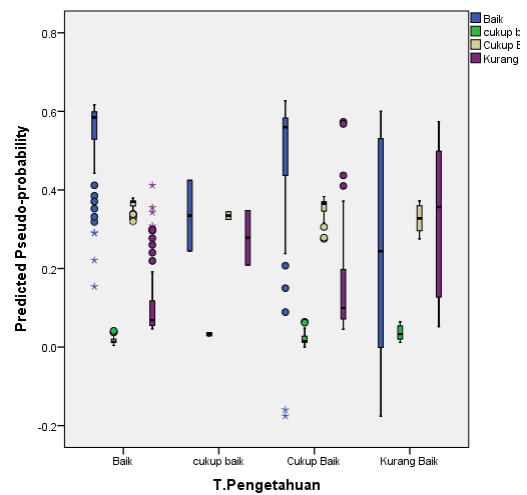
Gambar 1(a)



Gambar1(b)

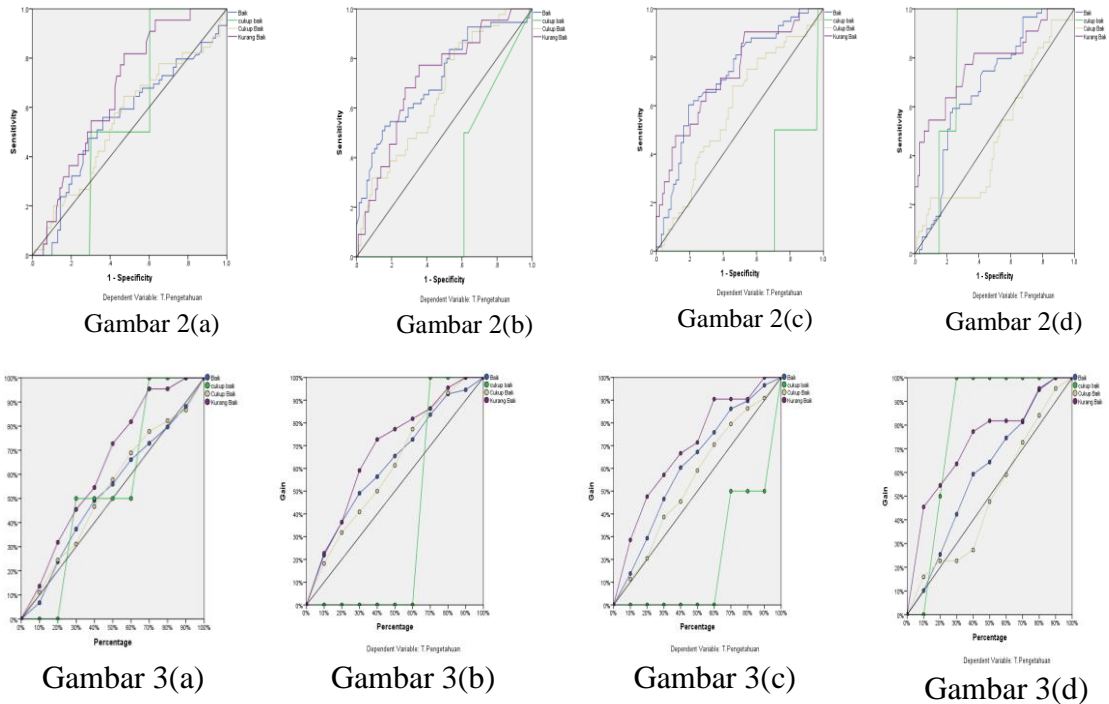


Gambar 1(c)



Gambar 1(d)

Gambar 1(a) dan 1(b) merupakan hasil klasifikasi dari jaringan arsitektur 2 hidden layer fungsi aktivasi *sigmoid* dan optimasi algoritma menggunakan *Conjugate gradient*, sedangkan Gambar 1(c) dan 1(d) merupakan hasil klasifikasi dari jaringan arsitektur 2 hidden layer fungsi aktivasi *Hyberbolic Tangent* dan optimasi algoritma menggunakan *Gradient descent*. Hasil terbaik diperoleh pada gambar 1(c) dengan tingkat akurasi prediksi training 55,9%, dan akurasi prediksi testing 46,9%.



### 3. Hasil dan Pembahasan

Dari hasil klasifikasi diperoleh hasil beragam dari input 2 *hidden layers*, dan penggunaan fungsi aktivasi sigmoid dan *Hyberbolic Tangent*. Hasil akurasi prediksi dapat dilihat pada table 2, sebagai berikut:

Tabel 2. Arsitektur jaringan

Hidden layers		Output layers	Optimization algorithm	Akurasi Prediksi	
Number of hidden layers	Activation Function	Activation Function		Training	Testing
2	Sigmoid	Sigmoid	Conjugate gradient	45,6%	48%
2	Hyberbolic Tangent	Hyberbolic Tangent	Conjugate gradient	50%	41,9%
2	Sigmoid	Sigmoid	Gradient descent	55,9%	46,9%
2	Hyberbolic Tangent	Hyberbolic Tangent	Gradient descent	53,1%	42,4%

Untuk data yang telah diuji dengan optimasi algoritma *Conjugate gradient* dan *Gradient descent*, hasil yang didapatkan dapat dilihat pada table 3, lihat tabel berikut:

Tabel 3. Hasil Perbandingan Algoritma Optimasi MLPNN

Optimization algorithm	Percent Incorrect Predictions	Sum of Squares Error	Training Time
Conjugate gradient	52%	8.106	0:00:00.06
Conjugate gradient	58,1%	9.489	0:00:00.09
Gradient descent	53.1%	10.253	0:00:00.06
Gradient descent	58.6%	9.354	0:00:00.07

### 4. Simpulan

Dari hasil analisis MLPNN, hasil yang ditunjukkan pada Tabel 2, yaitu diperoleh jumlah optimal unit tersembunyi untuk mendapatkan akurasi terbaik ditentukan berdasarkan *trial and error*, pada hasil

---

training MLPNN dengan menggunakan fungsi aktivasi output layer *sigmoid* hanya dapat menghasilkan akurasi klasifikasi 55,9% dan 46,9%. Dari hasil tersebut bahwa disimpulkan data yang digunakan dapat diterapkan menggunakan algoritma MLP neural network.

### Daftar Pustaka

- [1] WHO. Global Strategy for Infant and Young Child Feeding: The Optimal Duration of Exclusive Breastfeeding, 54th WHA. 2005.
- [2] Depkes. Strategi Nasional Peningkatan Pemberian ASI (PP-ASI). Jakarta : Departemen Kesehatan Republik Indonesia. 2001.
- [3] Irawati A, dkk. 2003. Pengaruh Status Gizi Selama Kehamilan dan Menyusui terhadap Keberhasilan Pemberian ASI. Penelitian Gizi dan Makanan (PGM); 26 (2): 10-19
- [4] Riset Kesehatan Dasar. Badan Penelitian dan Pengembangan Kesehatan Departemen Kesehatan, Republik Indonesia. 2010
- [5] Purnamawati, S. Faktor-Faktor yang Berhubungan dengan Pola Pemberian ASI pada Bayi Usia Empat Bulan (Analisis Data Susenas 2001). Badan Litbang Kesehatan.2003.
- [6] Marlenny, Finki Dona. "KLASIFIKASI TINGKAT PENGETAHUAN IBU DALAM PEMBERIAN ASI EKSLUSIF DENGAN METODE DECISION TREE." SEMNASTEKNOMEDIA ONLINE 5.1 (2017): 2-1.
- [7] Wu, Xindong, et al. "Top 10 algorithms in data mining." Knowledge and information systems 14.1 (2008): 1-37.
- [8] Witten, Ian H., et al. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.
- [9] Larose, Daniel T. Discovering knowledge in data: an introduction to data mining. John Wiley & Sons, 2014.
- [10] Han, Jiawei, Jian Pei, and Micheline Kamber. Data mining: concepts and techniques. Elsevier, 2011.
- [11] E. Alpaydin, Introduction to Machine Learning: Second Edition, London: The MIT Press, 2010.
- [12] R.-Z. Michal, B. Michael e K. and Ido, Learnability of periodic activation functions: General results, [http://www.ics.uci.edu/~michal/p3606\\_1.pdf](http://www.ics.uci.edu/~michal/p3606_1.pdf)., 1998.
- [13] W. L. K, "Receiver operating characteristic (ROC) analysis, Evaluating discriminance effects among decision support systems".