

# Clasificación en Python utilizando Azure como base de datos

Johnatan Alexander Mena Salcedo, Mateo Osorio Higueta

johnatanmenasalcedo@gmail.com

mateooh97@gmail.com

Instituto tecnológico Metropolitano

Medellín Colombia 2020

**Abstract**— Abstract Artificial intelligence will undoubtedly be the forerunner of the future of the technology of the future, where you can think of vehicles that ignite with recognition of physical features, interconnected cameras in a city that recognize the face of a criminal or computers that offer solutions without needing if programmed, understanding how the artificial intelligence works allows the development of a variety of programs that allow solutions to be provided based on autonomous learning, without a doubt a technological novelty unprecedented throughout history. [1]

## I. INTRODUCCION

### A. Antecedentes

En el año 1943 fue presentado el primer modelo de una neurona artificial por el psiquiatra Warren McCulloch y el matemático Walter Pitts, luego en la poca de 1958 fue creada por el investigador, psicólogo y filósofo Frank Rosenblatt quien al investigar el comportamiento de neurona diseñó un sistema matemático que permite emular este comportamiento, a este desarrollo lo llamó perceptrón, quien fue el precursor de los hardware [2], ordenadores o maquinas que emulan el comportamiento de una neurona, desarrollo inventado en la década de 1960 por Mark I. Finalmente en 1969 Minsky y Papert demostraron que el perceptrón simple tenía una limitación, la cual era no poder resolver problemas no lineales, esta limitación fue superada con la invención del perceptrón multicapa por Rumelhart entre otros en el año 1986. [1]

### B. Marco conceptual

Los algoritmos de inteligencia artificial son un diseño físico-matemático utilizado por ordenadores

para simular el comportamiento de una neurona o una red neuronal, todo esto con el fin de enseñarle a una maquina a partir de unos datos y unos resultados conocidos un fenómeno específico y con unos nuevos datos esperar unos resultados aproximados del programa anterior, en otras palabras, enseñarle a la maquina como aprender [3]. Como se menciona anteriormente Microsoft brinda una plataforma para almacenar proyectos destinados al manejo de datos, esta plataforma funciona adaptándose a varios lenguajes de programación y funciona de manera privada o pública [4].

### C. Importancia del tema

En la actualidad existen muchos softwares de programación diseñados para el procesamiento de datos y su clasificación por medio de la inteligencia artificial, en trabajos anteriores usamos el software Matlab como plataforma para clasificar predicción de fuerza a través de electroencefalografía y los resultados no fueron satisfactorios, esta vez se intentará trasladar la base de datos caracterizada al software Python y usaremos los mismos algoritmos de programación pero esta vez con un herramienta que permite almacenar datos en la nube y además permite su implementación desde otras plataformas, en este caso se trata de Microsoft Azure. [5]

## II. METODOLOGIA

Se extrae de Matlab la base de datos clasificada y etiquetada en formato. csv, posteriormente se carga esta base de datos en Python y se procede a cargarla en la base de datos creada en SQL server para posteriormente ser utilizada desde Azure, en

el software de programación Visual Code Estudios se trasladan los tres algoritmos de programación y se suben a la nube de Azure portal, desde esta plataforma se ejecutan los tres algoritmos de programación utilizando la base de datos cargada desde SQL server y se obtienen los siguientes resultados. [6] [7]

#### A. Resultados

| Nombre     | Tipo    | Tamaño  | Valor                            |
|------------|---------|---------|----------------------------------|
| AccActual  | float64 | 1       | 0.6114285714285714               |
| BestAcc    | float64 | 1       | 0.6136363636363636               |
| BestC      | int     | 1       | 0                                |
| BestGamma  | int     | 1       | 0                                |
| C          | float64 | (5,)    | [ 0.01 2.5075 5.005 7.5025 10. ] |
| Fn         | int64   | 1       | 53                               |
| Fold_Index | float64 | (3518,) | [4. 7. 8. ... 0. 2. 5.]          |
| Fp         | int64   | 1       | 83                               |

Figura 1. Resultados usando Naive Bayes

| Nombre     | Tipo    | Tamaño  | Valor                            |
|------------|---------|---------|----------------------------------|
| AccActual  | float64 | 1       | 0.53125                          |
| BestAcc    | float64 | 1       | 0.5738636363636364               |
| BestC      | int     | 1       | 0                                |
| BestGamma  | int     | 1       | 0                                |
| C          | float64 | (5,)    | [ 0.01 2.5075 5.005 7.5025 10. ] |
| Fn         | int64   | 1       | 174                              |
| Fold_Index | float64 | (3518,) | [6. 2. 0. ... 6. 0. 6.]          |
| Fp         | int64   | 1       | 0                                |

Figura 2. Resultados usando SVM

| Nombre     | Tipo    | Tamaño  | Valor                            |
|------------|---------|---------|----------------------------------|
| AccActual  | float64 | 1       | 0.5028571428571429               |
| BestAcc    | float64 | 1       | 0.5738636363636364               |
| BestC      | int     | 1       | 0                                |
| BestGamma  | int     | 1       | 0                                |
| C          | float64 | (5,)    | [ 0.01 2.5075 5.005 7.5025 10. ] |
| Fn         | int64   | 1       | 174                              |
| Fold_Index | float64 | (3518,) | [6. 2. 0. ... 6. 0. 6.]          |
| Fp         | int64   | 1       | 0                                |

Figura 3. Resultados usando KNN

### III. DISCUCION

Al analizar los resultados se observa que existe coincidencia en el desempeño de los algoritmos realizados tanto en Matlab como Python ya que el nivel de incertidumbre no es muy significativo, sin embargo, al comparar los tres algoritmos de clasificación se observa que el desempeño tanto en KNN como en SVM es similar y el mejor desempeño se octavo en el clasificador Naive Bayes, a continuación, se presenta una tabla comparativa. [8] [4]

| Clasificador | Acc   |
|--------------|-------|
| Knn vecinos  | 0.57  |
| SVM          | 0.57  |
| Naive Bayes  | 0.613 |

Tabla. 1 resultados clasificación

### IV. CONCLUSIONS

A pesar de usar de que usaron dos softwares de programación (Matlab y Python) para clasificóciar señales electroencefalográficas los resultados no variaron significativamente lo que significa que se debe implementar otra caracterización u otra forma de adquirir los datos. Por otra parte, ser resalta la importancia de contar con una base de datos en la nube ya que esto nos permite más accesibilidad y es un método seguro en casos de perdida de información importante, por lo que se resalta la importancia del software Azure función y las demás plataformas adyacentes que la complementan.

### REFERENCES

- [1] H. Chen, S. Yeh, K. Li, and C. Chen, "The movement-related cortical activities under the virtual environment with force feedback," in *Proc. Int. Conf. Platform Technology and Service*, 2015, pp. 11–12.
- [2] M. Akilli and N. Yilmaz, "Study of weak periodic signals in the EEG signals and their relationship with postsynaptic potentials," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 10, pp. 1918–1925, 2018.
- [3] P. Nielsen and U. Parui, *Microsoft SQL server 2008 bible*. John Wiley & Sons, 2011, vol. 607.
- [4] N. E. S. Alam, M. H. K. Azman, Z. M. Yusof, J. A. Shah, and K. Kadir, "Implementation of EEG approach for detection of upper limb movement," in *Proc. Measurement and Application (ICSIMA) 2019 IEEE Int. Conf. Smart Instrumentation*, 2019, pp. 1–5.
- [5] Z. Tang and J. MacLennan, *Data mining with SQL Server 2005*. John Wiley & Sons, 2005.
- [6] G. L. Calhoun and G. R. McMillan, "EEG-based control for human-computer interaction," in *Proc. Third Annual Symp. Human Interaction with Complex Systems. HICS'96*, 1996, pp. 4–9.
- [7] J. Anderson and J. Cho, "Software defined network based virtual machine placement in cloud systems," in *Proc. MILCOM 2017 - 2017 IEEE Military Communications Conf. (MILCOM)*, 2017, pp. 876–881.
- [8] B. Samal, A. K. Behera, and M. Panda, "Performance analysis of supervised machine learning techniques for sentiment analysis," in *Proc. Signal Processing and Security (ICSSS) 2017 Third Int. Conf. Sensing*, 2017, pp. 128–133.