Clasificación de dígitos manuscritos: una comparación entre KNN, SVM, Random Forests y Artificial Neural Networks (ANN).

Cavasin Nicolas - Universidad Nacional de Luján

**Resumen.** En el presente trabajo se documenta el rendimiento de la clasificación de dígitos manuscritos digitalizados al realizar una comparación entre algoritmos pertenecientes al aprendizaje automático supervisado que no formaron parte de la cursada de la asignatura.

Para ello, se utiliza un conocido dataset llamado MNIST

al aprendizaje automático supervisado haciendo uso de una red neuronal convolucional (CNN) para analizar canciones y responder ciertos interrogantes.

Para ello, se utilizó el dataset Free Music Archive[3], FMA de aquí en más, en su versión minimizada que posee un tamaño de 7.2GB y 8000 fragmentos de 30 segundos de duración de diferentes canciones.

El tipo de red neuronal seleccionado es un derivado de las redes neuronales artificiales (ANNs) y pertenece al subconjunto llamado Deep Learning, categoría que forma parte del amplio conjunto de algoritmos de aprendizaje automático agrupado bajo el concepto de Machine Learning.

Se expone brevemente cómo funcionan las CNN, se justifica el cómo y el porqué de la arquitectura adoptada, se documenta el proceso atravesado para su creación, entrenamiento y modelado y por último se muestran los resultados obtenidos.

1 Introducción

Las redes neuronales artificiales.

Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción. Esta es la introducción.

*Toda la información de contexto que permita entender el trabajo que se realizó. No más de una carilla.*

2 Materiales y Métodos

**2.1 - MNIST:**

El dataset utilizado es un derivado del dataset NIST

Toda la descripción del dataset y las técnicas que utilizaron, acá pueden utilizar tablas y gráficos que ayuden a explicar los datos pero no se incluye nada de resultados. Es solo el análisis exploratorio.

Acá también se describe todo el workflow de preprocesamientos y todas las transformaciones realizadas.

En está sección también se explican los experimentos que se realizaron. Ejemplo:

Se realizó una reducción de dimensionalidad con el método X y quedaron estas N variables, la variable objetivo utilizada fue la generada con el criterio de referencia dada en el enunciado del TP. Se corrió con esos datos un ajuste con J48 y se evaluó con la métrica Y.

**2.2 - PCA y t-SNE:**

Visualizing Datausingt-SNE => REF 8

How to uset t-SNE effectively => REF 9

Accelerating t-SNE using tree-based algorithms => REF 10

Visualizing high dimentional datasets using pca and t-SNE in python => REF 11

**2.3 - Normalización:**

Asdasdasdsa

**2.4 - K-fold:**

asdasdads

**2.5 - KNN:**

asdasdada

**2.6 - SVM:**

asdasd

**2.7 - Random forests:**

Asdasdasda

**2.8 - ANN:**

asdsada

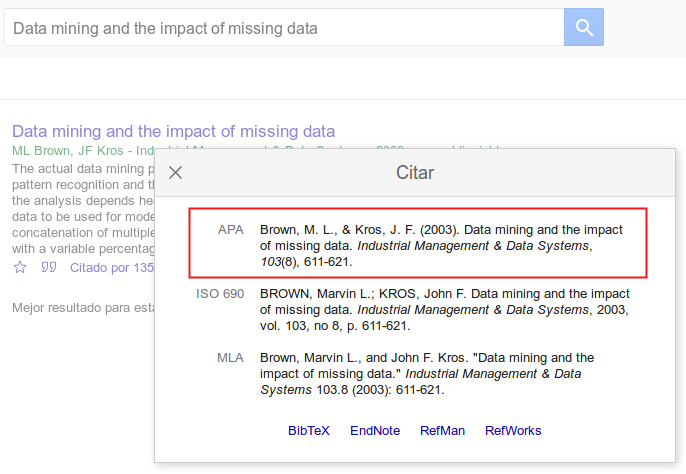
3 Resultados obtenidos

Los resultados de la configuración de preprocesamiento X con las N variables se muestran en la tabla XX.....

Referencias

1. Asd
2. Asd
3. Asd
4. Asd
5. Sad
6. Dsaasd
7. Sad
8. Maaten, L. V. D., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, *9*(Nov), 2579-2605.
9. <https://distill.pub/2016/misread-tsne/?_ga=2.135835192.888864733.1531353600-1779571267.1531353600>
10. Van Der Maaten, L. (2014). Accelerating t-SNE using tree-based algorithms. *The Journal of Machine Learning Research*, *15*(1), 3221-3245.
11. <https://towardsdatascience.com/visualising-high-dimensional-datasets-using-pca-and-t-sne-in-python-8ef87e7915b>

<https://towardsdatascience.com/musical-genre-classification-with-convolutional-neural-networks-ff04f9601a74>



Pueden obtenerlas ya formateadas las citas desde Scholar Google

|  |
| --- |
| **Aclaraciones Finales:**  La extensión final del trabajo no debe superar las 14 páginas con figuras y referencias.  En los casos que el trabajo sea un desarrollo de soporte a las decisiones el mismo debe estar disponible |