

April 29, 2024

## PROYECTO 1 — Redes Neuronales Artificiales y SVM

### 1. Red Neuronal Artificial (ANN)

#### 1.1. Optimización de Hiperparámetros y Frente de Pareto

En la Figura 1 se muestra cómo se relacionan los hiperparámetros a través de diversas pruebas de la optimización bayesiana de los hiperparámetros.

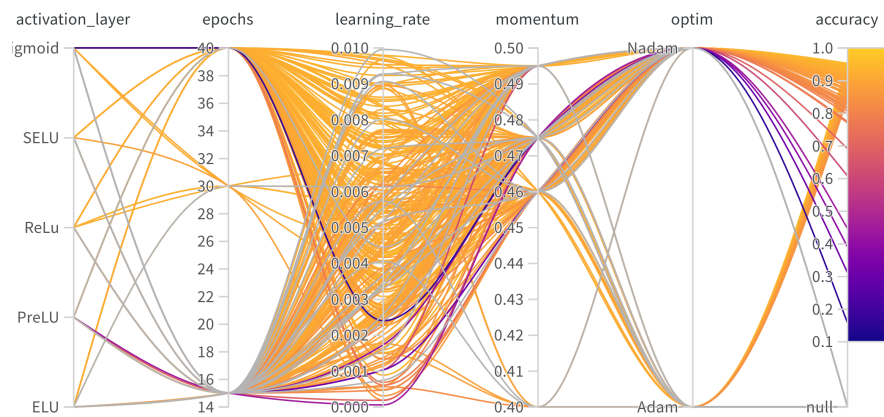


Figura 1: Relación entre los hiperparámetros y la exactitud (accuracy) del modelo ANN

En la Figura 2, se muestra el Frente de Pareto que ilustra la relación entre precisión y exhaustividad de los hiperparámetros identificados mediante la optimización bayesiana de hiperparámetros. Donde se obtiene que los puntos dominantes del Frente son los valores de la Tabla 1, para eso valores concretos **hidden\_size**, **learning\_rate**, **activation\_layer**, **optim** y **epochs**.

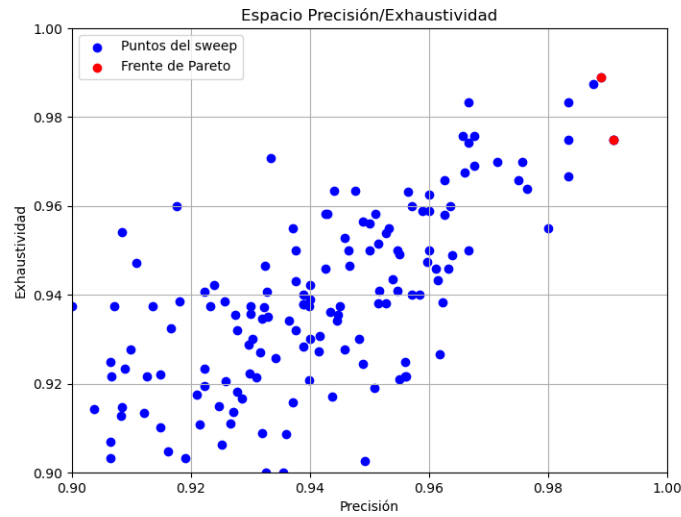


Figura 2: Frente de Pareto para la optimización de hiperparámetros del modelo ANN

Precisión (%)	Exhaustividad (%)	hidden_size	learning_rate	activation_layer	optim	epochs
98.89	98.89	[512, 256, 128, 64, 32]	0.0028718272	PreLU	Adam	15
99.09	97.5	[512, 256, 128, 64]	0.00382924790	PreLU	Adam	15

Tabla 1: Hiperparámetros óptimos del modelo ANN obtenidos con datos de prueba de MNIST

## 1.2. Evaluación del Modelo con Datos de Prueba de MNIST

En la Figura 3 presenta la matriz de confusión obtenida tras evaluar diferentes hiperparámetros y encontrar los valores óptimos.

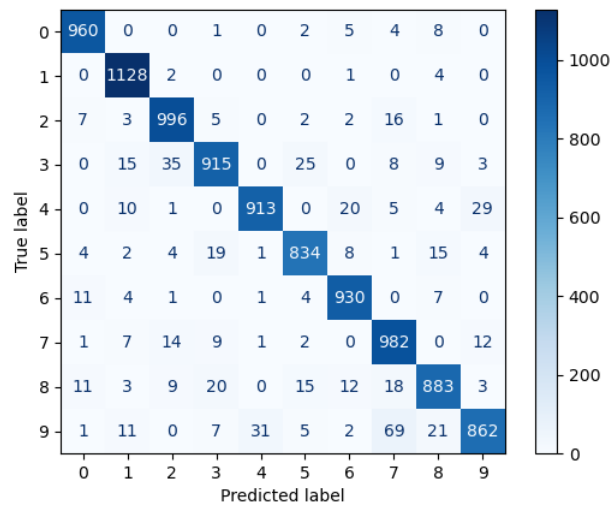


Figura 3: Matriz de confusión del modelo ANN con hiperparámetros optimizados y datos de prueba de MNIST

Y los resultados de precisión y exactitud correspondientes a la Figura 3 se detallan en la Tabla 2.

Precisión (%)	Exhaustividad (%)	Exactitud (%)
94.03	94.07	93.96

Tabla 2: Métricas de desempeño del modelo ANN con datos de prueba de MNIST

El desglose de cada una de estas clases y su precisión, exhaustividad, F1 y soporte se detallan en la Tabla 3

Tabla 3: Reporte de clasificación del modelo ANN con datos de prueba de MNIST

Clase	Precisión	Exhaustividad	Puntaje F1	Soporte
0	0.965	0.980	0.972	980
1	0.954	0.994	0.973	1135
2	0.938	0.965	0.951	1032
3	0.938	0.906	0.921	1010
4	0.964	0.930	0.947	982
5	0.938	0.935	0.937	892
6	0.949	0.971	0.960	958
7	0.890	0.955	0.922	1028
8	0.928	0.907	0.917	974
9	0.944	0.854	0.897	1009
<b>Accuracy</b>			0.940	10000
<b>Macro Avg</b>	0.941	0.940	0.940	10000
<b>Weighted Avg</b>	0.941	0.940	0.940	10000

### 1.3. Evaluación del Modelo con Set de Datos Grupal

En la Figura 4 se muestra la matriz de confusión obtenida tras evaluar el set de datos de prueba realizado por el grupo.

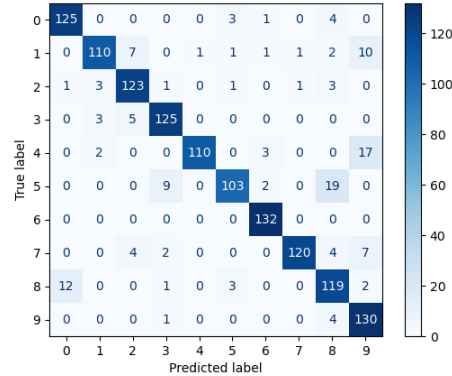


Figura 4: Matriz de confusión del modelo ANN con set de datos grupal

En la Tabla 4, se muestran los resultados de las métricas de desempeño obtenidos tras evaluar el set de datos de prueba realizado por el grupo.

Tabla 4: Métricas de desempeño del modelo ANN con set de datos grupal

Precisión (%)	Exhaustividad (%)	Exactitud (%)
89.46	90.25	89.47

Adicionalmente, en la Tabla 5 se presentan la precisión, exhaustividad, puntaje F1, y el soporte (cantidad de datos) para cada una de las clases.

Tabla 5: Reporte de clasificación del modelo ANN con set de datos grupal

Clase	Precisión	Exhaustividad	Puntaje F1	Soporte
0	0.906	0.940	0.923	133
1	0.932	0.827	0.876	133
2	0.885	0.925	0.904	133
3	0.899	0.940	0.919	133
4	0.991	0.833	0.905	132
5	0.928	0.774	0.844	133
6	0.950	1.000	0.974	132
7	0.984	0.876	0.927	137
8	0.768	0.869	0.815	137
9	0.783	0.963	0.864	135
<b>Accuracy</b>			0.895	1338
<b>Macro Avg</b>	0.903	0.895	0.895	1338
<b>Weighted Avg</b>	0.902	0.895	0.895	1338

## 1.4. Análisis de Resultados

En la Figura 1 se puede apreciar cómo para las distintas combinaciones de hiperparámetros es posible obtener distintos resultados en términos de su exactitud. Estos resultados son ocasionados no por algún dato en específico, sino por la combinación de determinados datos. Es decir, es la combinación entre una capa de activación específica, con una tasa de aprendizaje, un momentum y un optimizador puntuales quienes pueden generar mejores o peores exactitudes. En este caso, el peor caso corresponde a un *accuracy* (exactitud) de 0.157 (o 15.7 %), mientras que el mejor caso corresponde a una exactitud de 0.953 (95.3 %), y, tal y como se mencionó previamente, son ocasiones debido a combinaciones específicas de datos, que a grosso modo suelen brindar porcentajes de exactitud muy optimistas, siendo la mayoría de ellos siendo superiores al 0.8 (u 80 %).

Por otro lado, la Figura 2 permite observar los datos que son posteriores resumidos en la Tabla 2, donde se observa que el método Adam, con una capa de activación PreLU y una tasa de aprendizaje muy pequeña permiten obtener buenos resultados de precisión y exhaustividad. Se concluye, a partir de la misma tabla, que una mayor cantidad de capas permite una mayor exhaustividad, mientras que menos capas aseguran la precisión. Esto muy posiblemente ocasionado por un sobreajuste que se genera a mayor cantidad de capas, por lo que el modelo compromete su precisión pero genera una mayor exhaustividad.

Estos últimos porcentajes tienen su similitud con los obtenidos en la Tabla 1, en donde se puede observar que los números suelen ser correctamente predichos. Cabe aclarar que, este porcentaje difiere a los expuestos en la Tabla 2, debido a que con cada ejecución del código se carga un conjunto distinto (*batch*) de la base de datos de MNIST, por lo que estos casos pueden prometer ciertos márgenes de exactitud, mas no un número exacto.

Finalmente, el modelo evaluado en el set de datos grupal consiguió un gran porcentaje de exactitud, tal y como presenta la Figura 4. Adicionalmente, la Tabla 4 detalla los resultados obtenidos en términos de precisión, exhaustividad y exactitud. Es posible apreciar como estos tres resultados son (levemente) inferiores a los conseguidos con MNIST, lo cual era de esperar debido a la manera tan distinta que tienen los dígitos de ser dibujados por los miembros del grupo en relación a los disponibles en MNIST. Esto ocasionaba un conjunto de datos retardador y que por tanto supondría la posibilidad de enfrentarse a una exactitud un poco más baja que para el conjunto inicial.

## 2. Máquina de Soporte Vectorial (SVM)

### 2.1. Optimización de Hiperparámetros y Frente de Pareto

En la Figura 5 se muestra cómo se relacionan los hiperparámetros a través de diversas pruebas de la optimización bayesiana de los hiperparámetros.

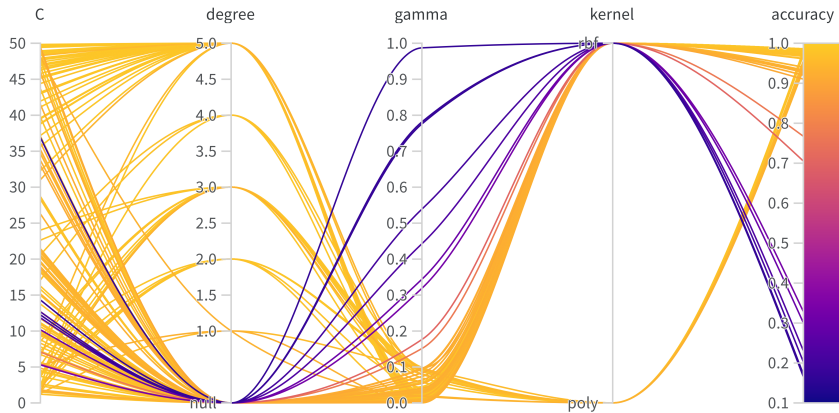


Figura 5: Relación entre los hiperparámetros y la exactitud (accuracy) del modelo SVM

En la Figura 6, se muestra el Frente de Pareto que ilustra la relación entre precisión y exhaustividad de los hiperparámetros identificados mediante la optimización bayesiana de hiperparámetros. Donde se obtiene que los puntos dominantes del Frente son los valores de la Tabla 6, para eso valores concretos **C**, **Kernel** y **Gamma**.

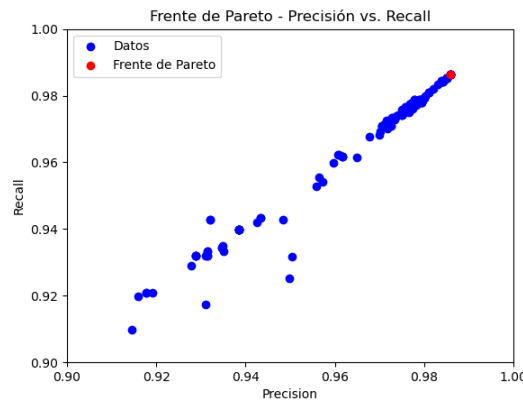


Figura 6: Frente de Pareto para la optimización de hiperparámetros del modelo SVM

Precisión (%)	Exhaustividad (%)	C	Kernel	Gamma
98.5975	98.6358	42.98620007644219	rbf	0.021332766431507475

Tabla 6: Hiperparámetros óptimos del modelo SVM obtenidos con datos de prueba de MNIST

## 2.2. Evaluación del Modelo con Datos de Prueba de MNIST

En la Figura 7 presenta la matriz de confusión obtenida tras evaluar diferentes hiperparámetros y encontrar los valores óptimos.

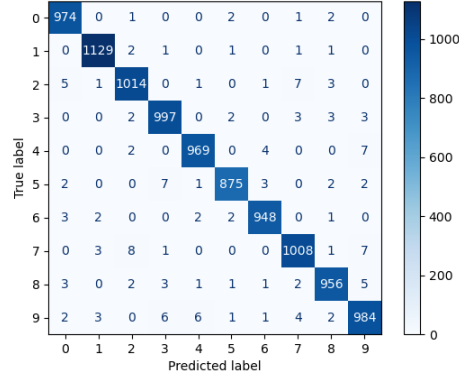


Figura 7: Matriz de confusión del modelo SVM con hiperparámetros optimizados y datos de prueba de MNIST

En la Figura 6, se ilustra el Frente de Pareto, que destaca la relación entre precisión y exhaustividad de los hiperparámetros obtenidos mediante la optimización bayesiana. Los puntos dominantes en el Frente corresponden a los valores específicos de **C**, **Kernel** y **Gamma** presentados en la Tabla 7.

Precisión (%)	Exhaustividad (%)	Exactitud (%)
98.54	98.53	98.54

Tabla 7: Métricas de desempeño del modelo SVM con datos de prueba de MNIST

Tabla 8: Reporte de clasificación del modelo SVM con datos de prueba de MNIST

Clase	Precisión	Exhaustividad	Puntaje F1	Soporte
0	0.985	0.994	0.989	980
1	0.992	0.995	0.993	1135
2	0.984	0.983	0.983	1032
3	0.982	0.987	0.985	1010
4	0.989	0.987	0.988	982
5	0.990	0.981	0.985	892
6	0.990	0.990	0.990	958
7	0.982	0.981	0.981	1028
8	0.985	0.982	0.983	974
9	0.976	0.975	0.976	1009
<b>Accuracy</b>			0.985	10000
<b>Macro Avg</b>	0.985	0.985	0.985	10000
<b>Weighted Avg</b>	0.985	0.985	0.985	10000

## 2.3. Evaluación del Modelo con Set de Datos Grupal

En la Figura 8 se muestra la matriz de confusión obtenida tras evaluar el set de datos de prueba realizado por el grupo.

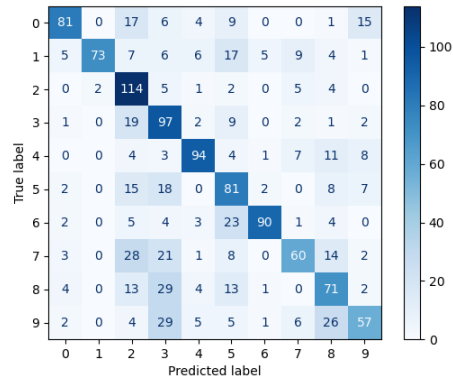


Figura 8: Matriz de confusión del modelo SVM con set de datos grupal

En la Tabla 9, se muestran los resultados de las métricas de desempeño obtenidos tras evaluar el set de datos de prueba realizado por el grupo. Adicionalmente, en la Tabla 10 se presentan la precisión, exhaustividad, puntaje F1, y el soporte (cantidad de datos) para cada una de las clases.

Tabla 9: Métricas de desempeño del modelo SVM con set de datos grupal

Precisión (%)	Exhaustividad (%)	Exactitud (%)
66.56	61.26	61.14

Tabla 10: Reporte de clasificación del modelo SVM con set de datos grupal

Clase	Precisión	Exhaustividad	Puntaje F1	Soporte
0	0.810	0.609	0.695	133
1	0.973	0.549	0.702	133
2	0.502	0.850	0.631	133
3	0.445	0.729	0.553	133
4	0.783	0.712	0.746	132
5	0.474	0.609	0.533	133
6	0.900	0.682	0.776	132
7	0.659	0.438	0.526	137
8	0.493	0.518	0.505	137
9	0.606	0.422	0.498	135
<b>Accuracy</b>		0.611		1338
<b>Macro Avg</b>	0.665	0.612	0.617	1338
<b>Weighted Avg</b>	0.664	0.611	0.616	1338

## 2.4. Análisis de Resultados

Como parte de los resultados obtenidos de la evaluación del modelo con los datos de prueba MNIST muestran se obtiene en la Figura 7 la matriz de confusión proporciona una comprensión más profunda del rendimiento del modelo en términos de clasificaciones correctas e incorrectas. La Tabla 7 muestra los resultados tras la evaluación de diferentes hiperparámetros lo cual condujo a la identificación de valores óptimos que maximizaron la exactitud del modelo a un 98.54 %. Se evaluaron dos tipos de kernels, polinomial y base radial, con el objetivo de mejorar la exactitud del modelo. A pesar de que ambos kernels fueron evaluados exhaustivamente, se determinó que el kernel rbf proporcionó el mayor nivel de exactitud en comparación con el kernel polinomial. La alta precisión y exhaustividad indican que el modelo es capaz de clasificar correctamente la mayoría de las instancias positivas, minimizando tanto los falsos positivos como los falsos negativos. Además, la alta exactitud refleja la capacidad general del modelo para clasificar correctamente las mayoría de instancias.

Analizando diferentes combinaciones de hiperparámetros para el modelo SVM y generamos el frente de Pareto, que representa las soluciones que no pueden ser mejoradas en términos de precisión o exhaustividad que que maximiza tanto la precisión como la exhaustividad mediante optimización bayesiana. En la Tabla 6 se muestran estos hiperparámetros representan una configuración óptima que equilibra la capacidad de generalización del modelo con su capacidad para identificar positivos verdaderos. En la Figura 5 se muestran las diferentes combinaciones de hiperparametros y en la Figura 6 el punto en el Frente de Pareto.

Al evaluar el modelo SVM con el set de datos grupal, se observó una disminución significativa en los porcentajes de precisión, exhaustividad y exactitud en comparación con el desempeño anterior con el conjunto de datos MNIST. Esto era de esperar, dado que el set de datos de dígitos dibujados por los miembros del grupo es distinto al conjunto de datos MNIST utilizado anteriormente. En la Tabla 9 se detallan las métricas, se muestra una precisión del 66.56 %. La naturaleza diferente de los dígitos dibujados a mano puede introducir variaciones y complejidades que el modelo previamente entrenado no ha aprendido a manejar adecuadamente.

Además se observa que en la Tabla 9 el número que peor precisión y puntaje F1 tiene es el 9, por lo que es el dato que más se confunde, y según la Figura 8 este confunde principalmente con el 3 y el 8. Este es un comportamiento inesperado ya que se esperaba que el número que se confundiera fuera el 4 o 7, debido a su similitud de escritura. El segundo peor predicción de número es el 8, este se confunde con el 3, este número es esperable que se confunda con el 8 porque un 3 es la mitad de un 8.

También se realiza la observación que los datos con mejores precisión son 2 y 3. Se realiza una hipótesis que la razón de la alta confusión de los números 8 y 9, son principalmente a que un sobreajuste del modelos por lo datos entrenamiento de MNIST.



## Referencias

[1] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Online Publisher, 2012.

**Adicionalmente** se debe de destacar la utilidad de herramientas de inteligencia artificial utilizadas como apoyo a lo largo de este documento, como lo fue [ChatGPT](#).

*Submitted by David Badilla Marin - Lizzy González Alvarado - Ignacio Grané Rojas - Carlos Andrés Mata Calderón on April 29, 2024.*