



EXÁMEN UNIDAD II.

Reconocimiento de Firmas con red neuronal backpropagation (NNBP).

Asignatura:

“Inteligencia Artificial”

Maestro:

Francisco Javier Luna Rosas

Alumnos:

- José Alfredo Díaz Robledo
- Luis Pablo Esparza Terrones
- Luis Manuel Flores Jiménez
- Juan Francisco Gallo Ramírez

***Ingeniería en Computación
Inteligente
3er Semestre***



Introducción

El proceso de verificación de firmas es una técnica fundamental para determinar si una firma adicional corresponde a la misma persona. Esta verificación desempeña un papel crucial en la autenticación de la personalidad. Se puede realizar tanto en línea como fuera de línea, dependiendo de si se captura la firma en tiempo real o se utiliza una imagen de una firma previamente realizada en papel. En ambos casos, se extraen una serie de características de la firma que permiten diferenciar entre firmas genuinas y falsificadas, incluso las falsificaciones hábiles. Estas características se obtienen a partir de múltiples ejemplares de la firma, lo que forma un conjunto de patrones utilizado para entrenar y probar un clasificador.

La verificación en línea implica el uso de dispositivos como plumas instrumentadas o tabletas digitalizadoras para capturar tanto la forma como el movimiento dinámico de la mano del firmante, lo que resulta en una verificación más precisa debido a la información dinámica disponible. Por otro lado, la verificación fuera de línea se aplica a firmas que han sido registradas como imágenes estáticas, lo que significa que se pierde valiosa información dinámica en el proceso.

Este proceso de verificación de firmas se divide en cuatro etapas: adquisición de datos de cheques con firmas, preprocesamiento y extracción de firmas, almacenamiento y transformación de datos relevantes, y aprendizaje supervisado a través de técnicas de clasificación, como redes neuronales de retropropagación. La implementación de estas etapas es fundamental para el éxito de la verificación de firmas fuera de línea, ya que la falta de información dinámica hace que sea un desafío lograr un rendimiento tan alto como en la verificación en línea.

Este trabajo aborda estos pasos en detalle, destacando la importancia de cada uno de ellos en el proceso de verificación de firmas. Además, se proporciona una visión general de la arquitectura utilizada para el reconocimiento de patrones y se establece un marco para la implementación práctica.

Etapa 1.

Base de Datos de Cheques y Algoritmo de Segmentación.

La base datos de cheques se adjuntan en el envío de la entrega, sin embargo, el link para acceder a ella es el siguiente:

- [Cheques](#)

El algoritmo de segmentación de la firma fue entregado en tiempo y forma en la fecha correspondiente de entrega, aquí se encuentra el link hacia el recurso:

- [ExtractorFirma.zip](#)

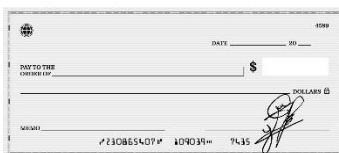
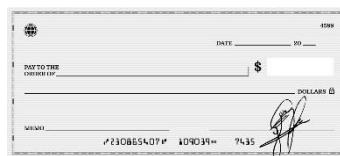
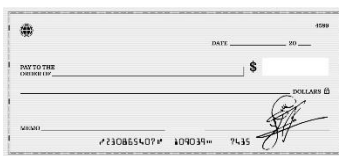
Etapa 2.

Base de Datos de Firmas.

La base datos de firmas extraídas de los cheques se adjuntan en el envío de la entrega, sin embargo, el link para acceder a ella es el siguiente:

- [Firmas](#)

Resultados:





Etapas 3.

Dataset de patrones de firmas.

El dataset de las firmas se incluye en el algoritmo de la red neuronal que más adelante se analizará. Este consta de 15 patrones reales, 50 patrones sintéticos positivos y 60 patrones sintéticos negativos.

Los **patrones reales** se obtuvieron mediante los elementos estructurantes que obtuvimos de una firma base, para después intersecar estos elementos con los

elementos corrientes de la firma a analizar posteriormente se hizo el conteo de los elementos y de esta forma se estableció el dataset.

Los **patrones sintéticos positivos** son los generados a partir de los patrones reales, se aplica una fórmula para que tengan valores aleatorios, pero simulando patrones reales. La fórmula es la siguiente:

$$\frac{1}{n} \sum_{k=1}^n x_k \pm \sigma$$

Los **patrones sintéticos negativos** son los generados aleatoriamente, estos patrones no tienen ninguna relación con los patrones reales y por ello simularán una firma distinta.

El total de elementos fue de 125, 15 patrones reales, 50 patrones sintéticos positivos y 60 patrones sintéticos negativos.

Dataset:

E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20	Valor
0	0	0	6	219	537	210	0	0	0	6	153	0	0	0	0	0	0	0	0	1
75	732	675	789	741	1122	543	0	0	0	309	435	150	918	150	93	0	0	0	0	1
0	0	0	0	18	477	0	0	0	0	240	114	0	159	0	0	0	0	0	0	1
0	0	111	174	0	69	408	0	0	0	0	63	0	210	9	69	0	0	0	0	1
0	210	0	0	294	276	459	0	0	0	0	84	21	255	18	0	0	0	0	0	1
12	0	6	369	609	291	144	0	0	0	0	87	108	606	0	0	0	0	0	0	1
0	141	0	0	192	630	135	0	0	0	192	333	0	0	0	0	0	0	0	0	1
0	123	0	0	246	627	318	0	0	0	0	39	51	0	0	0	0	0	0	0	1
0	105	0	0	0	972	210	0	0	0	0	57	54	0	0	0	0	0	0	0	1
75	384	405	225	150	510	510	0	0	0	0	21	0	0	0	0	0	0	0	0	1
72	252	15	450	3	216	24	0	0	0	300	177	150	234	24	0	0	0	0	0	1
0	0	0	0	543	474	174	0	0	0	0	0	75	159	0	0	0	0	0	0	1
0	156	75	0	237	384	483	0	0	0	42	201	66	714	114	60	0	0	0	0	1
0	0	0	396	30	555	342	0	0	0	267	129	0	282	0	0	0	0	0	0	1
24	0	0	0	0	873	378	0	0	0	0	0	120	255	150	63	0	0	0	0	1
...

En la tabla anterior solo se muestran los 15 patrones reales, el dataset completo se encuentra en el siguiente link:

- [DataSetFirmas.xlsx](#)



Etapa 4.

Aprendizaje supervisado en el reconocimiento de patrones de la firma off-line. Máquinas de Aprendizaje Supervisado (Clasificación/Predicción).

Análisis:

1. Encabezado y metadatos:

En las primeras líneas del código, se incluyen comentarios con información relevante sobre el propósito del programa, la institución, la materia, el examen, los nombres de los miembros del equipo, el nombre del profesor y la fecha.

2. Importación de bibliotecas:

El código comienza importando diversas bibliotecas, lo que sugiere que se utilizarán en el programa. Estas bibliotecas incluyen string, random, numpy, PIL (Pillow), openpyxl, pandas, IPython.display, sklearn, y más. Estas bibliotecas son esenciales para diversas tareas, como procesamiento de imágenes, manipulación de datos y aprendizaje automático.


3. Definición de variables:

Se definen varias variables importantes, como el ancho y largo de las imágenes, la cantidad de elementos estructurantes y las dimensiones de estos elementos.

4. Carga de la firma base:

Se carga una imagen de una firma base y se redimensiona a las dimensiones especificadas.

5. Extracción de elementos estructurantes:



La imagen de la firma base se divide en varios elementos estructurantes. Estos elementos se recortan y almacenan en una lista.

6. Conversión de elementos estructurantes en matrices:

Los elementos estructurantes se convierten en matrices binarias para su procesamiento posterior. Se realiza una binarización de las imágenes.

7. Creación de un archivo Excel:

Se crea un archivo Excel que servirá para almacenar los datos de las firmas. Se establecen los encabezados para las columnas.

8. Generación de patrones reales:

Se extraen patrones reales de la base de datos de 15 firmas. Estos patrones se comparan con los elementos estructurantes y se almacenan en el archivo Excel. También se etiquetan como "1" (positivos) en la columna correspondiente.

9. Generación de patrones sintéticos positivos:

Se generan patrones sintéticos positivos basados en los patrones reales. Estos patrones tienen valores aleatorios, pero simulan firmas reales. Se etiquetan como "1" en la columna correspondiente.


10. Generación de patrones sintéticos negativos:

Se generan patrones sintéticos negativos aleatorios que no tienen relación con las firmas reales. Estos patrones se etiquetan como "0" en la columna correspondiente.

11. Guardado del archivo del dataset:

Se guarda el archivo Excel que contiene los datos de las firmas y sus etiquetas.

12. Preparación para el entrenamiento de la red neuronal:



Se lee el archivo Excel, se separan las matrices categorizadas y predictorias, y se dividen en conjuntos de entrenamiento y prueba.

13.Creación de una red neuronal:

Se crea una instancia de una red neuronal utilizando la biblioteca scikit-learn. El solver utilizado es 'lbfgs', y se ajusta a los datos de entrenamiento.

14.Predicciones y evaluación del modelo:

Se hacen predicciones en los datos de prueba y se calculan índices de calidad del modelo, como la matriz de confusión, precisión y error globales.

Implementación y Evaluación:

La implementación de esta actividad se realizó mediante el uso de Jupyter Notebooks en lenguaje Python.

El algoritmo tomó muchas líneas de código para mostrarlas en el presente archivo sin que se vea con el orden correspondiente, por lo que es posible visualizarlo mediante el siguiente link.

El vínculo al archivo se muestra a continuación:

- [ReconocimientoFirma.html](#)

Gracias al uso de la herramienta de Jupyter Notebooks se realizó la evaluación en el mismo documento que se muestra anteriormente.



Conclusión

En resumen, el trabajo se centra en el proceso de verificación de firmas y se divide en cuatro etapas: adquisición de datos de cheques con firmas, preprocesamiento y extracción de firmas, creación de un dataset de patrones de firmas y aprendizaje supervisado utilizando una red neuronal.

El proceso comienza en la etapa 1 con la adquisición de datos de cheques con firmas, que pueden realizarse en tiempo real o utilizando imágenes de firmas previamente realizadas en papel. A continuación, se extraen características de las firmas, lo que permite diferenciar entre firmas genuinas y falsificadas. El conjunto de patrones se utiliza para entrenar y probar un clasificador.

En la etapa 2, se crea una base de datos de firmas a partir de los cheques. Los resultados de esta etapa se proporcionan a través de enlaces para acceder a los datos.

En la etapa 3, se crea un dataset que consta de patrones reales, patrones sintéticos positivos y patrones sintéticos negativos. Los patrones reales se obtienen de la firma base, y los patrones sintéticos se generan tanto a partir de los patrones reales como de manera aleatoria. Se muestra una parte del dataset en la conclusión, pero el conjunto completo está disponible en un enlace.

Finalmente, en la etapa 4, se utiliza una red neuronal para el aprendizaje supervisado en el reconocimiento de patrones de firmas. Se describen los pasos de entrenamiento del modelo y se evalúa su rendimiento utilizando medidas como la matriz de confusión, precisión y error global.

En general, el trabajo se enfoca en la implementación práctica de un sistema de verificación de firmas fuera de línea y resalta la importancia de cada etapa en el proceso. Los enlaces a los recursos y datos utilizados están disponibles para acceder a la información completa.



Referencias

MathWorks. (s.f.). *Tipos de operaciones morfológicas*. Obtenido de https://es.mathworks.com/help/images/morphological-dilation-and-erosion.html#mw_870854ff-1cfa-4a13-94a5-58fd427d9d55