

TRABAJO PRÁCTICO FIS SUGENO Y MAMDANI

Integrantes:

Bianchini, Pilar.

de Aza, Paloma.



INTRODUCCIÓN:

Uno de los contenidos que abarca la materia, es el de Lógica Difusa (Fuzzy Logic). Según Michael Negnevitsky, en su libro “Artificial Intelligence: A Guide to Intelligent Systems”, podemos definir a la Lógica Difusa como “un conjunto de principios matemáticos para la representación del conocimiento, basados en grados de pertenencia en lugar de una pertenencia estricta o binaria propia de la lógica clásica.” En otras palabras, a diferencia de la lógica booleana en la cual tenemos dos resultados posibles únicamente (0 o 1, verdadero o falso), la lógica difusa es multivaluada. Se ocupa de grados de pertenencia y grados de verdad. La lógica difusa utiliza el continuo de valores lógicos entre 0 (completamente falso) y 1 (completamente verdadero). En lugar de solo “blanco y negro”, emplea el espectro de colores, aceptando que las cosas pueden ser parcialmente verdaderas y parcialmente falsas al mismo tiempo. Para ello, empleamos un Modelo de Inferencia, el cual, en base a una serie de reglas (compuestas por un antecedente y consecuente), funciones de pertenencia, y datos de entrada, es decir, en base a una serie de datos “difusos”, podemos inferir un valor de verdad, que esté más cerca de una representación en el mundo real.

En el presente trabajo, utilizaremos Sistemas de Inferencia Difusa (FIS) propuestos por Sugeno y Mamdani para estudiar el funcionamiento de los mismos, y lograr mapear datos de entrada difusos en datos de salida (valores objetivo), y lograr inferir.....???

Se presentará una metodología similar entre las actividades propuestas, y a continuación se mostrarán los resultados obtenidos para las diferentes actividades.

ACTIVIDAD 1a- Sugeno VDA

METODOLOGIA:

En los sistemas de inferencia difusa de Sugeno, una de las características principales es que a partir de la generación de reglas IF-THEN, compuestas por un antecedente y consecuente, se puede modelar y predecir un grado de pertenencia, para situaciones en las cuales definir un resultado, a partir de datos “difusos” no es una cuestión “binaria” (es decir, asignar 1 o 0). Para ello, es importante la generación de los antecedentes y consecuentes en estas reglas. Los antecedentes se conforman a partir de los diferentes datos (variables) de entrada que disponemos. Mientras que los consecuentes se conforman como funciones matemáticas de las variables de entrada X_1, X_2, \dots, X_n , y coeficientes a_1, a_2, \dots, a_n , determinados durante el modelado. De este modo, cuando se cumple el antecedente de una regla, el consecuente calcula una salida numérica basada en las entradas.

Tanto para la resolución de la actividad 1a y 1b, en la cual usamos el sistema de inferencia de Sugeno, utilizamos en su mayoría una implementación provista por la cátedra, perteneciente al Ing. Daniel Albornoz. La misma emplea un algoritmo de clustering sustractivo para la determinación de centros de cluster, lo cual, como indica el sistema, a su vez permite hallar la función matemática (de forma gaussiana) correspondiente a una determinada regla IF-THEN.



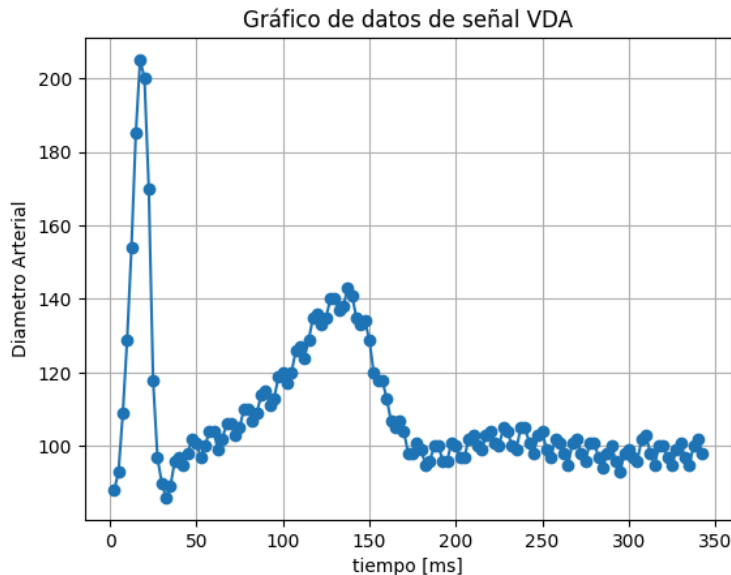
En la implementación dada, cobraba importancia tanto la elección de los hiperparámetros de los algoritmos, como la elección de los datos de entrada. Cabe aclarar que tanto en la actividad 1a como 1b, de los conjuntos de datos provistos por la cátedra, se tomaron 2 variables de entrada para caso, para poder así obtener una representación en un solo plano, de manera que pueda observarse en un gráfico.

Uno de los hiperparámetros que cobraba importancia a la hora de entrenar el modelo, era la elección de un Radio de vecindad del clustering sustractivo, el cual, al variar, aumenta o disminuye la cantidad de clusters, por lo tanto, la cantidad de reglas halladas. Al variar esto, también varía esta “clasificación” que se les puede dar a los datos de entrada, lo cual también influye en el nivel de especificidad o certeza del resultado, para que al momento de inferir un resultado, se obtenga un nivel de alto nivel de coherencia. Este Radio de vecindad, también está fuertemente ligado con la cercanía que tienen los datos en los datasets propuestos, por ende, un radio pequeño o grande, puede afectar a la generación de clusters.

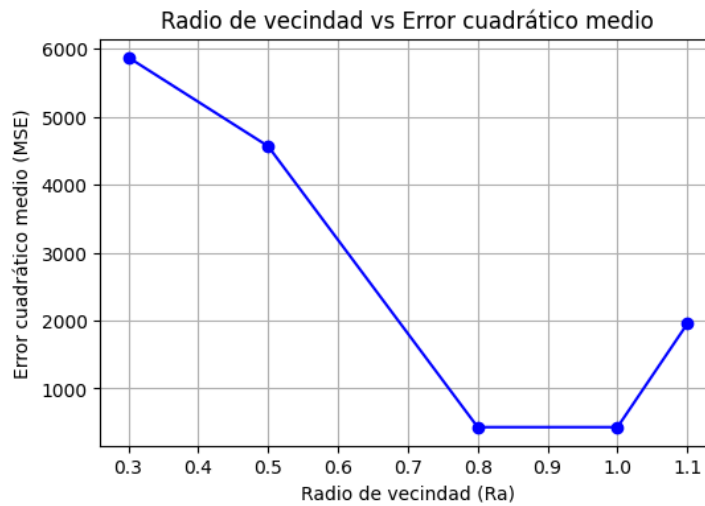
A partir de estos radios de vecindad, se entrena el modelo (uno para cada radio), y se obtiene, a partir de la evaluación del mismo, el error cuadrático medio. Esto nos ayuda a ver qué modelo, a partir de qué hiperparámetros, obtuvo el menor error, y es el que usamos para realizar un sobremuestreo, para inferir datos que concuerden con el dato objetivo y con lo esperado.

RESULTADOS:

1. Gráfico de los datos de una de las señales de Variación de Diámetro Arterial.



2. Variando el hiperparámetro Radio de Vecindad entre clusters, y entrenando una serie de modelos diferentes, obtuvimos los siguientes resultados:

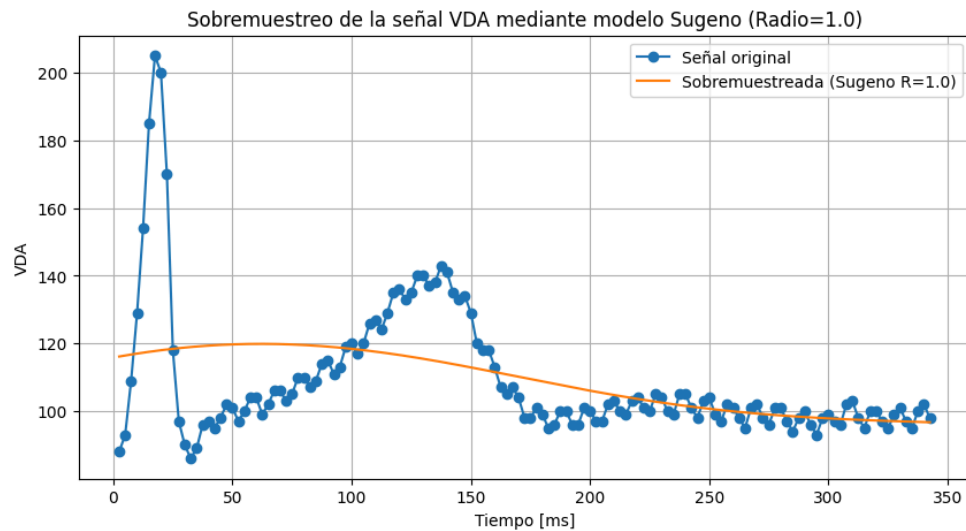


Además vemos la cantidad de reglas que se obtuvo para cada modelo FIS Sugeno:

Radio de vecindad	Cantidad de Reglas	MSE Obtenido
0.3	13	5875.88643705
0.5	9	4567.59624383
0.8	5	429.976762292
1.0	4	429.93387749
1.1	3	1953.0589421

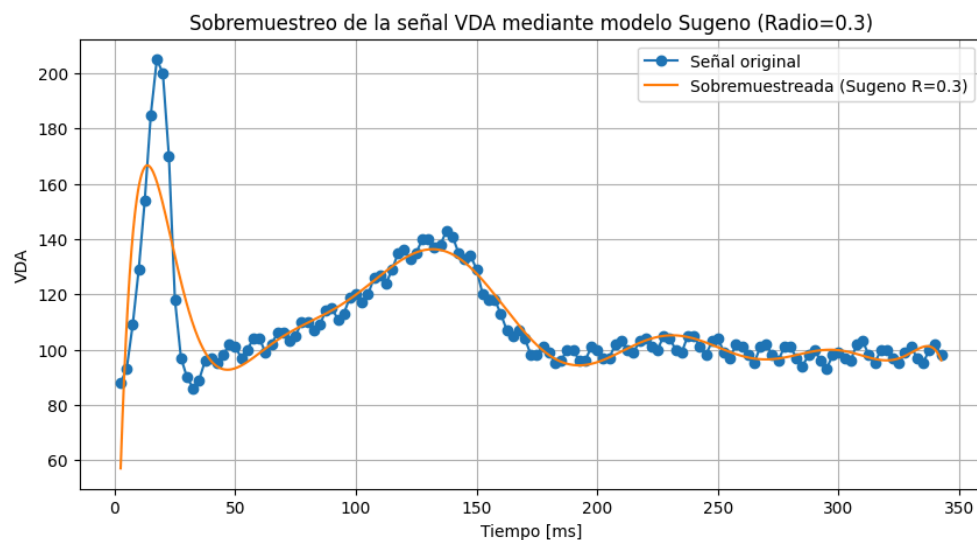
- En vista de que el mejor modelo obtenido es aquel en el que el radio de vecindad es 1.0, ese es el que elegiremos para sobremuestrear.
- Con el modelo seleccionado y entrenado de la actividad anterior. Sobremuestreamos la señal y obtenemos

Con Radio de vecindad=1.0



Sin embargo, vimos que la curva estaba lejos de acercarse a la señal original y decidimos probar con otro modelo, a pesar no ser el de menor error cuadrático medio.

Con Radio de vecindad=0.3



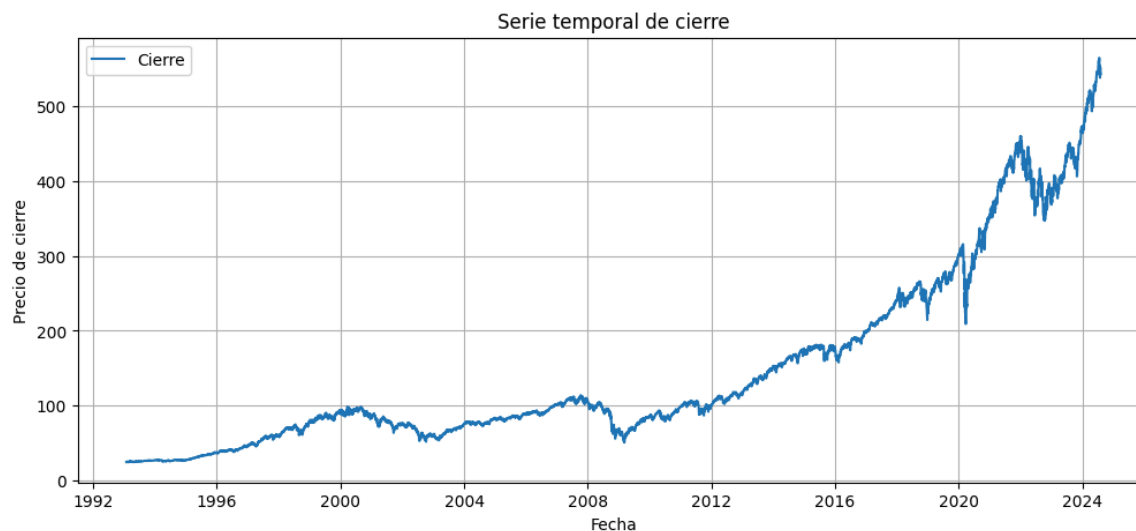
Generamos una escala en el tiempo más densa, para poder interpolar la curva. En vista de la primera señal obtenida, viendo como esta se aleja de la señal inicial, decidimos disminuir el Radio de vecindad, ya que de esa forma, obteniendo una mayor cantidad de reglas, (por ende de funciones) para clasificar los datos de entrada, y a pesar de tener un error cuadrático medio mayor, se obtiene una señal que parece asemejarse más a la inicial.

ACTIVIDAD 1b- Sugeno- SPY

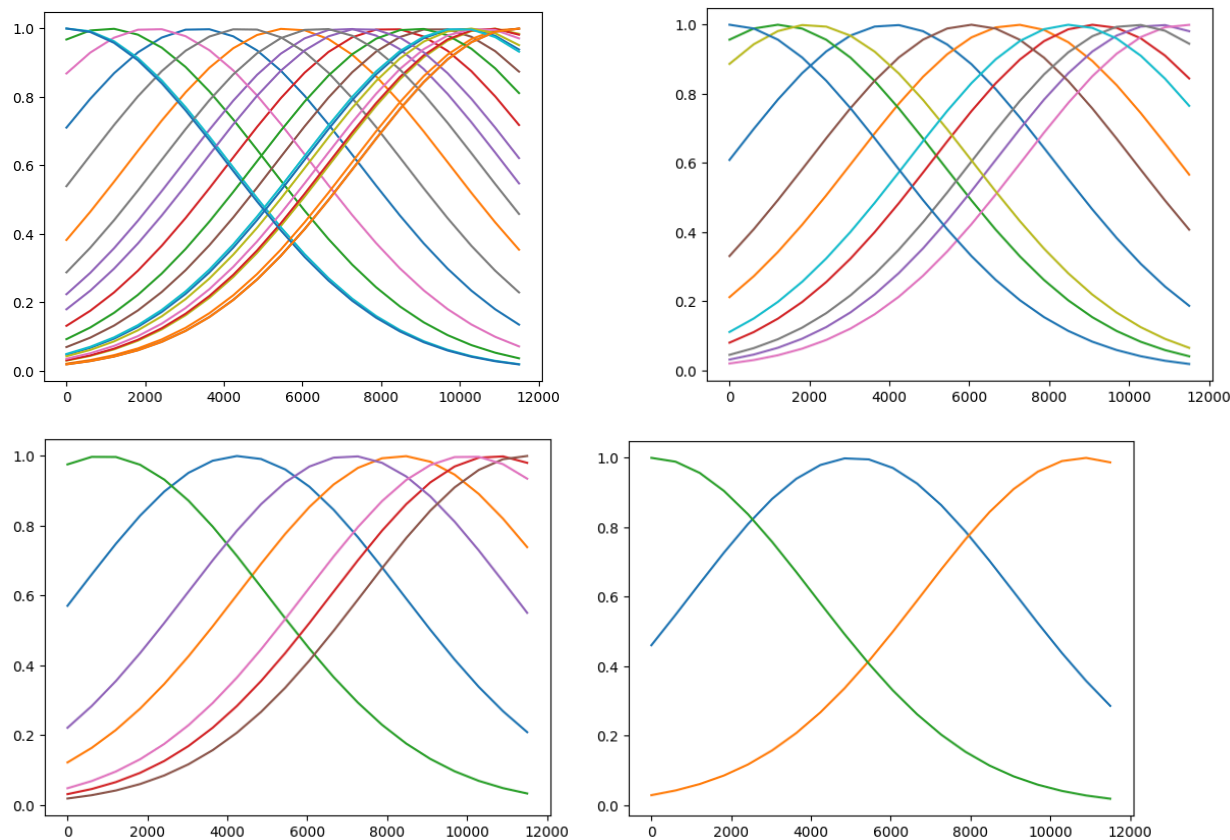
1.Gráfico de los datos:

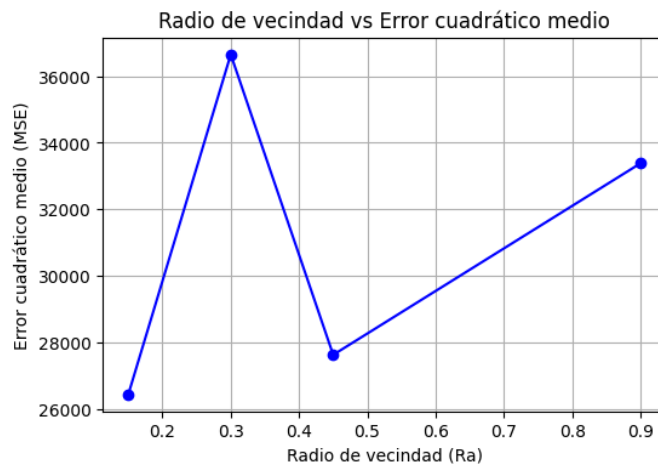


Es importante aclarar que: si bien no podemos graficar la totalidad de los datos al tener 11 posibles tipos de datos de entrada, decidimos graficar el tiempo vs. el precio de cierre. Para graficar el tiempo, se debió pasar del formato fecha, a un dato numérico para poder trabajar con el modelo, lo cual, podría llegar a generar inconsistencias futuras a la hora de inferir sobre datos sobre otros valores de tiempo.



2. Con $\text{radios}=[0.15,0.3,0.45,0.9]$ (de izquierda a derecha de arriba a abajo)



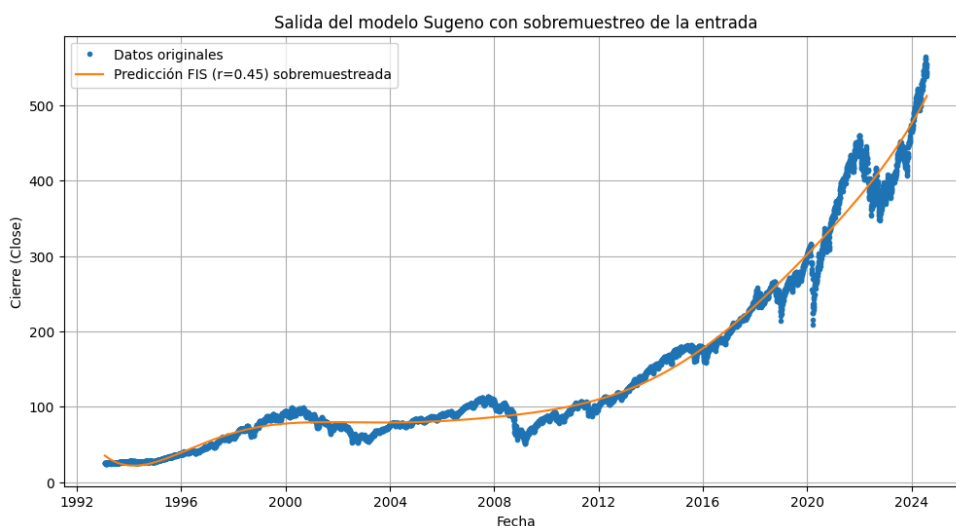


Además vemos la cantidad de reglas que se obtuvo para cada modelo FIS Sugeno:

Radio de vecindad	Cantidad de reglas	MSE Obtenido
0.15	22	26439.62174718
0.3	11	36650.21518918
0.45	7	27626.70552436
0.9	3	33380.06140551

3. En vista de que el mejor modelo obtenido es aquel en el que el radio de vecindad es 0.45, ese es el que elegiremos para sobremuestrear. Tenemos en cuenta que al tener un dataset tan grande, es posible querer una mayor cantidad de clusters para su clasificación.

4.

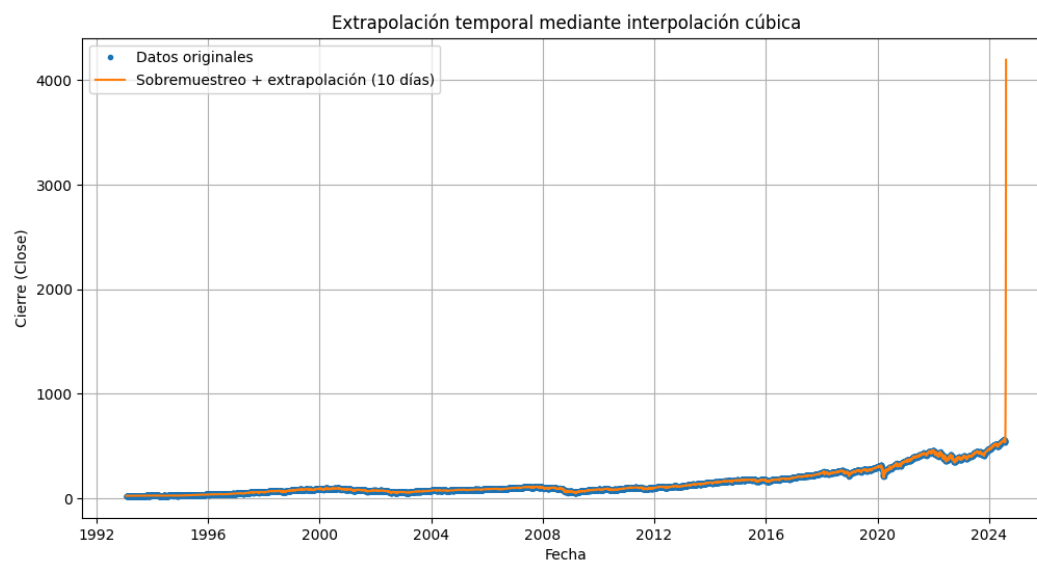




Usar el modelo difuso entrenado con Radio de vecindad=0.45 para evaluar muchos más puntos de entrada (sobremuestreados) y observar cómo el modelo reconstruye la función continua de salida. El sobremuestreo densifica la entrada para lograr una señal continua. El resultado indica como el modelo “aprendió” la relación entre los datos de entrada para lograr inferir una salida, aunque no del toda precisa, coherente con los datos iniciales.

Tal vez una forma de ajustarse todavía más a los datos originales, sea generando una mayor cantidad de clusters (reglas), con un radio de vecindad más pequeño, ajustándose al ruido local.

5. Queremos lograr extender el eje temporal más allá del último valor real del dataset, para “imaginar” cómo seguiría la curva del Cierre si los días continuaran.



La forma resultante del modelo difuso puede ser más suave o más ruidosa según el radio de vecindad utilizado, porque este parámetro controla cuántas reglas se generan al agrupar los datos: un radio grande (por ejemplo, 0.45–1.0) produce menos reglas y, por lo tanto, un comportamiento más general y una curva más suave; mientras que un radio pequeño (0.2–0.3) crea más reglas y un ajuste más detallado, capturando incluso pequeñas variaciones o ruido local. Usamos interpolación para extrapolar los datos, porque el conjunto de datos está definido en pasos diarios (con algunos días faltantes) y representa una serie temporal continua. Esto nos permite estimar valores intermedios y prolongar la tendencia más allá del rango observado de manera coherente con la estructura temporal del dataset, a pesar de que se observa que para inferir sobre datos futuros, comienzan a mostrarse inconsistencias esperables.

ACTIVIDAD 2b-Mamdani - Docentes:

METODOLOGIA:



Se diseñó un Sistema de Inferencia Difuso tipo Mamdani para estimar la nota final de un alumno a partir de dos entradas:

- Nota de examen: variable numérica continua en el rango $[0,100]$. Representa el rendimiento objetivo del estudiante. Tiene mayor peso en el resultado del estudiante.
- Nota de concepto: variable lingüística con tres valores posibles: regular, bueno, excelente. Esta variable modula la nota del examen, pudiendo incrementarla o disminuirla levemente.

La salida del sistema es la nota final, expresada en una escala de 1 a 10.

Metodología del profesor.

- Con las notas bajas y media bajas se es muy exigente: el concepto negativo resta mucho y el concepto bien o excelente suman muy poco a la nota
 - Las notas medias se consideran muy sensibles, ya que según esta nota el alumno puede aprobar o desaprobado: el concepto regular descuenta moderadamente, el concepto bien aumenta moderadamente y el concepto excelente suma mucho.
 - El profesor considera que las notas medias altas y altas son poco relevantes en el cambio de nota, pero de ellas se espera que el concepto sea bueno. Por lo tanto un concepto regular resta mucho y el concepto bueno y excelente suma poco.
 - Por último, las notas muy altas son muy estimadas por el profesor, por lo tanto el concepto regular resta muy poco y el concepto bien y excelente suman mucho.
-
- Términos difusos de nota: baja, media baja, media, media alta, alta y muy alta
 - Términos difusos de cambio de nota: negativo, negativo moderado, positivo moderado, positivo.
 - Reglas:
1. If Nota Baja and Concepto Regular then Ajuste Negativo.
 2. If Nota Baja and Concepto Bueno then Ajuste Positivo Moderado.
 3. If Nota Baja and Concepto Excelente then Ajuste Positivo Moderado.
 4. If Nota Media Baja and Concepto Regular then Ajuste Negativo.
 5. If Nota Media Baja and Concepto Bueno then Ajuste Positivo Moderado.
 6. If Nota Media Baja and Concepto Excelente then Ajuste Positivo Moderado.
 7. If Nota Media and Concepto Regular then Ajuste Negativo Moderado.
 8. If Nota Media and Concepto Bueno then Ajuste Positivo Moderado.
 9. If Nota Media and Concepto Excelente then Ajuste Positivo.
 10. If Nota Media Alta and Concepto Regular then Ajuste Negativo.

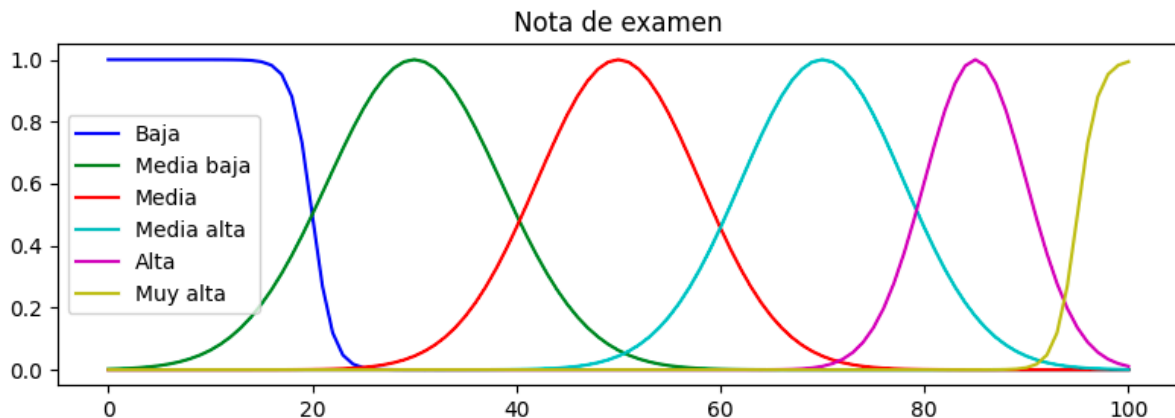


11. If Nota Media Alta and Concepto Bueno then Ajuste Positivo Moderado.
12. If Nota Media Alta and Concepto Excelente then Ajuste Positivo Moderado.
13. If Nota Alta and Concepto Regular then Ajuste Negativo.
14. If Nota Alta and Concepto Bueno then Ajuste Positivo Moderado.
15. If Nota Alta and Concepto Excelente then Ajuste Positivo Moderado.
16. If Nota Muy Alta and Concepto Regular then Ajuste Negativo Moderado.
17. If Nota Muy Alta and Concepto Bueno then Ajuste Positivo.
18. If Nota Muy Alta and Concepto Excelente then Ajuste Positivo.

RESULTADOS:

Funciones de pertenencia

Nota examen:



- Baja y Muy Alta: función sigmoidea.
- Media Baja, Media, Media Alta y Alta: función gaussiana.

Las categorías extremas (notas muy bajas o muy altas) presentan una menor ambigüedad, ya que resulta sencillo identificar cuándo una nota es claramente insuficiente o sobresaliente. Por eso se utilizaron funciones sigmoideas, que marcan límites más definidos en los extremos.

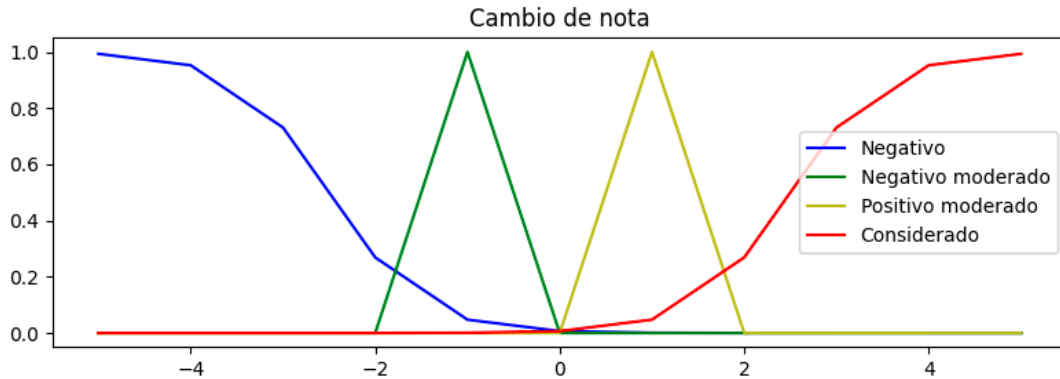
En cambio, las calificaciones intermedias se caracterizan por una mayor incertidumbre: no existe una frontera nítida entre lo que puede considerarse una nota media, media baja o media alta.

Concepto:



Todas funciones discretas de un valor. Máxima pertenencia para un valor y 0 para el resto de valores

Cambio de la nota:



- Función de pertenencia Negativo y Considerado: Sigmoidea
- Función de pertenencia Negativo moderado y Positivo moderado: Triangular

Para una nota de examen y un concepto bueno, se obtiene una nota final de 5.7.

CONCLUSIÓN:

Se desarrolló un sistema difuso de tipo Mamdani capaz de calcular la nota final de un alumno a partir de su desempeño en el examen y su nota de concepto. El modelo adopta un enfoque alternativo que estima el incremento o ajuste sobre la nota del examen, en lugar de calcular directamente la nota final. Esta estrategia permite evitar errores en los valores extremos y mantener un equilibrio adecuado entre las entradas, otorgando peso al examen y utilizando el concepto únicamente como modulador.

El resultado es una solución precisa y de implementación sencilla, que combina la objetividad del rendimiento numérico con la valoración cualitativa del docente. De este modo, el examen conserva el peso principal en la calificación, mientras que el concepto aporta una modificación razonable y acotada.

En cuanto a los modelos FIS Sugeno, podemos decir que los resultados obtenidos fueron coherentes con lo esperado, siempre considerando la naturaleza de los datos con los que trabajamos, ya que esta puede generar salidas que difieran ligeramente de lo previsto. Si bien, como mencionamos, se experimentó con la cantidad de reglas generadas para cada modelo a partir de la variación del Radio de vecindad —empleado como hiperparámetro en la función de clustering sustractivo—, también quedó pendiente analizar posibles cambios si se hubieran ajustado otros hiperparámetros, como el Acceptance Ratio o el Reject Ratio.

Por otro lado, es interesante observar cómo la variación del Radio de vecindad y su relación con el error cuadrático medio (MSE) obtenido para cada modelo pierde relevancia al



sobremuestrear la señal de entrada, tal como se aprecia en la primera actividad, aunque no ocurre lo mismo en la segunda.