

# Trabajo técnico: Seguros de autos y propensión a comprar.

Ricardo Yahir German Perez

En este trabajo se usan datos de una empresa de seguros, que contiene la información sobre el cliente de dicha compañía anónima, por temas de seguridad. En esta base de datos podemos encontrar un total de 86 variables que incluyen datos de uso de los productos de la compañía y dato sociodemográficos que se derivan en códigos postales.

## AVISO LEGAL

Este conjunto de datos es propiedad y es suministrado por la compañía holandesa de minería de datos Sentient Machine Research, y está basado en datos comerciales del mundo real. Se permite el uso de este conjunto de datos y la información adjunta solo para fines de investigación y educación no comercial.

## DATOS:

Esta base de datos está dividida en dos partes: el conjunto de entrenamiento y de evaluación, cada observación corresponde a un código postal. Las variables que empiezan con M se refieren a estadísticas sociodemográficos del código postal, mientras que las variables que comienzan con P y A se refieren a la propiedad de productos y estadísticas de seguros en el código postal.

La variable objetivo que se emplea en este trabajo es la de  $Y = \text{APERSAUT}$ , así como variables predictoras que serán las personas que se encuentran en un determinado grupo social, las cuales las variables eran:

- MSKA: Clase social A
- MSKB1: Clase social B1
- MSKB2: Clase social B2
- MSKC: Clase social C
- MSKD: Clase social D
- MOPLHOOG: Nivel de educación alto
- MOPLMIDD: Nivel de educación medio
- MOPLLAAG: Nivel de educación bajo

- MBERZELF: Emprendedor
- MBERBOER: Agricultor
- MBERMIDD: Gestión intermedia
- MBERARBG: Obreros cualificados
- MBERARBO: Obreros no cualificados

Cada uno de estos datos estará dividido en varios subconjuntos, para ser específicos en 3, los datos por nivel de educación donde se incluirán las variables nivel de educación (MOPLHOOG MOPLMIDD MOPLLAAG), por clases sociales (MSKA, MSKB1, MSKB2, MSKC, MSKD) y por tipo de empleo (MBERZELF, MBERBOER, MBERMIDD, MBERARBG, MBERARBO), esto con la intención de que se haga las probabilidades bayesianas y extraer cual es la más relevante de ella y con base a ello hacer un análisis, también lo mediremos con una red neuronal, para ver que modelo es mejor.

### **TEOREMA DE BAYES:**

El teorema de bayes es un pilar de las estadísticas y de la probabilidad, ya que esta nos explica la posibilidad de que pase un evento ante una serie de eventos o de características que pasan a su alrededor o en conjunto, en el caso del seguro y de las variables seleccionadas, como es que se da la adquisición de seguros ante las distintas clases sociales en las que se encuentran los clientes. Es por ello por lo que en este trabajo se usara el teorema de bayes, dado que le daremos un enfoque distinto a lo que se da en economía, que es un enfoque frecuentista. “la importación del teorema de bayes radica en que nos permite actualizar nuestras creencias a medida que obtenemos nueva información” <sup>1</sup>(Olguin A, 2024, p.1) de esta manera nos muestra que se puede expresar de la siguiente manera:

$$P(A|B) = P(B|A) * P(A) / P(B)$$

Donde:

- $P(A|B)$  es la probabilidad condicional de que ocurra A dado que B ha ocurrido. Esto se interpreta como la probabilidad posterior de A dado que se observó B.
- $P(B|A)$  es la probabilidad de observar B dado que A ha ocurrido, y se llama verosimilitud

---

<sup>1</sup> Olguin, Alfredo. 2024. «Una explicación introductoria al teorema de Bayes». Diapositivas. Facultad de Economía UNAM. 5 de septiembre de 2024.

- $P(A)$  es la probabilidad marginal de A, conocida como la probabilidad a priori de A, antes de observar B.
- $P(B)$  es la probabilidad marginal de B, también llamada la evidencia o el factor de normalización

Para este trabajo, establecemos los parámetros correspondientes con la demostración anterior donde

- $P(A|B)$ . La probabilidad de que el cliente haya comprado un seguro, dado que pertenece a una determinada clase social
- $P(B|A)$  La probabilidad de que el cliente pertenezca a una determinada clase social dado que compro un seguro de autos.
- $P(A)$  La probabilidad simple de tener un seguro de autos
- $P(B)$  La probabilidad simple de pertenecer a alguna de las clases sociales

Nota: se hará esto con los 3 subconjuntos para determinar las variables más representativas de los subconjuntos y hacer una probabilidad condicional de seguros dado ciertas variables seleccionadas. Por ejemplo

- $P(\text{Seguro} \mid \text{Clase social, Educación, Trabajo})$ . La probabilidad de que el cliente haya comprado un seguro, dado que pertenece a una determinada clase social, determinado tipo de empleo y determinado nivel de empleo.

Dado este planteamiento anterior nos hará más eficiente determinar en qué clase social, nivel de educación y tipo de empleo es la que hace la mayor adquisición de seguros de autos, para ellos se empleó el uso de lenguajes de programación como Python para filtrar y hacer el manejo de los datos de una manera más sencilla, se usaran bibliotecas como pandas, para la manipulación y manejo de datos, también Scikit-learn para el aprendizaje automatizado.

### **Redes Neuronales:**

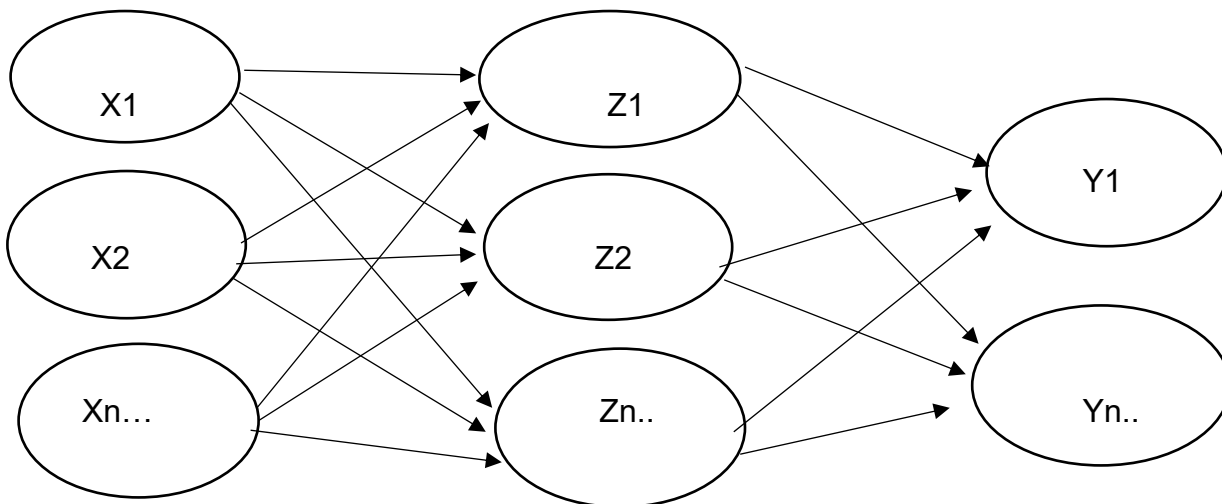
Con la implementación de las nuevas tecnologías, podemos hacer trabajos que normalmente son muy robustos o que pueden llegar a ser muy tardados de una manera mucho más eficiente, no solo en trabajos que son manuales, sino en cualquier tipo de trabajos, con las redes neuronales, podemos llegar a ser mucho más eficiente en la toma de decisiones para

cualquier área, ya sea de salud, de negocios, de educación , entre otros, para no caer en un error es importante saber cómo es que estas funcionan, aunque es un poco complejo como funcionan las redes neuronales, trataremos de explicar cómo funcionan de una manera simple.

Izaurieta (s.f) Las actividades de investigación desarrolladas en torno al estudio de redes neuronales artificiales, simplemente redes neuronales o neuroredes, están motivadas en modelar la forma de procesamiento de la información en sistemas nerviosos biológicos. Especialmente, por la forma de funcionamiento del cerebro humano, que es completamente distinta al funcionamiento de un computador digital. <sup>2</sup>

Si bien el proceso de una red neuronal no es algo que sea nuevo, pero si es algo que ha experimentado un gran cambio en los últimos días, por la cantidad de información que podemos procesar de forma simultánea, a diferencia de los computadores que son en forma secuencial, pueden hacer un proceso mucho más eficiente y de forma rápida, pero este proceso no siempre es precisamente bueno en todos los aspectos, ya que un proceso sencillo, como lo es el teorema de bayes, puede resultar mucho mejor que la implementación de las redes neuronales.

Una red neuronal empieza con una serie de datos que serán recibidos para el procesamiento, estos serán los datos de entrada, después pasaran por un proceso de capas para determinar los pesos de cada variable de entrada, esta es una capa oculta que realiza cálculos para saber estos pesos, esta capa es necesaria para hacer una función de activación para los próximos datos de entrada y finaliza con datos de salida generando datos predictivos.



---

<sup>2</sup> Izaurieta, Fernando. s. f. «Redes Neuronales Artificiales». *Universidad de Concepción*.

En el esquema anterior las X serán nuestros datos de entrada, todos estos datos son las que recoge de nuestra muestra, cada uno de estos datos pasan en una capa oculta que es la Z donde se harán una serie de operación para asignarle un peso y así una función de activación. La función de activación es aquella que permita que la red pueda aprender haciendo un enlace entre los datos de entrada y salida quitando la linealidad del modelo. Por último, pasarán a nuestros datos de salida que serán las Y, aunque este es un ejemplo burdo, pero es sencillo de entender cómo es que funciona de manera simple. El siguiente código en Python nos muestra cómo es que se hace esta red neuronal después de procesar los datos y limpiarlos

Imagen 1: Elaboración de una red neuronal con Python

```
#RED NEORONAL
# conjunto de entrenamiento y prueba
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
                                                    random_state=42)

# Escalar datos
escalar = StandardScaler()
X_train_scaled = escalar.fit_transform(X_train)
X_test_scaled = escalar.transform(X_test)

# Construir y entrenar la red neuronal
modelo = MLPClassifier(hidden_layer_sizes=(64, 32), max_iter=1000,
                        random_state=42, verbose=True)
modelo.fit(X_train_scaled, y_train)

# Precision y evaluacion del modelo
y_pred = modelo.predict(X_test_scaled)
print("Precision:", accuracy_score(y_test, y_pred))

# permutaciones alatorias
importancia = permutation_importance(modelo, X_test_scaled, y_test,
                                     n_repeats=10, random_state=42)
```

Como podemos ver, primero pasa por un conjunto de entrenamiento donde obtiene todos los datos de entrada, aquí el 80% de los datos serán para entrenar y en 20% para prueba. Pasa por un escalar para estandarizar los datos y así obtener un mejor rendimiento de los datos. En la construcción de la red neuronal será necesaria asignarles pesos en las capas ocultas ya mencionadas, después se evaluará en modelo y a la vez obtener la precisión de este, por último, el modelo hará que permutaciones aleatorias para medir la precisión de este modelo.

El código anterior fue obtenido de la documentación de [scikit-learn.org](https://scikit-learn.org)<sup>3</sup>, donde se pueden encontrar distintos ejemplos de redes neuronales para cualquier fenómeno que se guste estudiar.

### **Probabilidades Simples (Probabilidad de seguros de autos, nivel de educación, tipo de trabajo) :**

La probabilidad clásica es aquella en la que un suceso en un experimento aleatorio es posible que sucede, en este en nuestro trabajo,

Valdez y Alfaro (2018). Sea  $n(S)$  es el número de elementos, igualmente posibles y mutuamente excluyentes, del espacio muestral  $S$  de un experimento aleatorio, y sea  $n(A)$  el número de elementos de un evento cualquiera  $A$  de ese espacio muestral.

La probabilidad de que ocurra el evento  $A$ , al realizar el experimento, es la proporción de  $n(A)$  con respecto a  $n(S)$ .<sup>4</sup>

Por lo que cada uno de los subconjuntos de variables se seleccionaran a aquellos a los que tengas un determinado empleo, pertenezcan a determinada clase social y determinado nivel de educación, por lo que a cada subconjunto se le hará la probabilidad condicional y vemos que tenemos lo siguientes datos

Nivel de educación	Tipo de empleo	Clase social	Seguros de autos
MOPLLAAG: 0.948643	MBERMIDD: 0.885435	MSKC: 0.937479	APERSAUT: 0.511336310
MOPLMIDD: 0.927345	MBERARBO: 0.833734	MSKB2: 0.829955	
MOPLHOOG: 0.631226	MBERARBG: 0.799553	MSKB1: 0.767606	
	MBERZELF: 0.283580	MSKA: 0.701477	
	MBERBOER: 0.282721	MSKD: 0.552216	

<sup>3</sup> «1.17. Neural Network Models (Supervised)». s. f. Scikit-Learn. [https://scikit-learn.org/1.5/modules/neural\\_networks\\_supervised.html](https://scikit-learn.org/1.5/modules/neural_networks_supervised.html).

<sup>4</sup> Valdez y Alfaro, Irene Patricia. 2018. «PROBABILIDAD Y ESTADÍSTICA». Diapositivas. DVB. junio de 2018. [http://www.dcb.unam.mx/profesores/irene/Notas/Tema\\_2-1\\_pantcompl.pdf](http://www.dcb.unam.mx/profesores/irene/Notas/Tema_2-1_pantcompl.pdf).

Como se puede observar, la variable más representativa por la probabilidad clásica o simple, por nivel de educación es MOPLLA2"AG con un 0.948643 que representa el nivel alto de educación de educación, en el tipo de empleo es de MBERMIDD 0.885435 que es gestión intermedia, en la clase social es de MSKC 0.937479 que es la clase social C, en otras palabras, la probabilidad de extraer un dato de estas variables y que si pertenezca a una clase social C, que sea de gestión intermedia y que tenga educación alta. En cuanto a la variable objetivo que es del 0.5113363105462041, lo que quiere decir es que si extraigo un dato al azar voy a tener un 51% de que tenga un seguro de autos.

Aunque este no es un dato tan relevante, nos da la tentativa sobre las probabilidades condiciones, ya que para sacar estas probabilidades condicionales se requiere de las probabilidades simples o clásicas, como se mostró hace un momento. Recordemos se pueden calificar en dos tipos de resultados Valdez y Alfaro (2018). Nos menciona que "pueden ser determinístico es aquel cuyos resultados se pueden predecir de antemano y los probabilístico que es aquel en el que, para las limitaciones actuales del conocimiento científico, no se puede predecir con certeza el resultado" <sup>5</sup>en ese caso se usan los probabilísticos.

### **Probabilidades condicionales, P (Seguro | Clase social, Educación, Trabajo):**

Primero definimos que es una probabilidad condicional según Valdez y Alfaro (2018) " $P(A|B)$  y se lee la probabilidad de A dado B.  $P(A|B)$  equivale a calcular la probabilidad de A cuando el espacio muestral se reduce a B", <sup>6</sup>como ya lo habíamos definidos es una probabilidad que depende de un espacio muestral, de probabilidades simples tanto del evento de la variable de respuesta y explicativas por ello que se usa la verosimilitud

La verosimilitud es la que permite realizar inferencias acerca de su valor a partir de un conjunto de observaciones, en el caso anterior tenemos que sacar estos datos para hacer las probabilidades condicionales, aunque no es necesario mostrarlas si no más bien como es que obtiene este

---

<sup>5</sup> Valdez y Alfaro, Irene Patricia. 2018. «PROBABILIDAD Y ESTADÍSTICA». Diapositivas. DVB. junio de 2018. [http://www.dcb.unam.mx/profesores/irene/Notas/Tema\\_2-1\\_pantcompl.pdf](http://www.dcb.unam.mx/profesores/irene/Notas/Tema_2-1_pantcompl.pdf).

<sup>6</sup> Valdez y Alfaro, Irene Patricia. 2018. «PROBABILIDAD Y ESTADÍSTICA». Diapositivas. DVB. junio de 2018. [http://www.dcb.unam.mx/profesores/irene/Notas/Tema\\_2-1\\_pantcompl.pdf](http://www.dcb.unam.mx/profesores/irene/Notas/Tema_2-1_pantcompl.pdf).

dato, debido a esto un ejemplo, supongamos que tenemos datos de ciertas variables, para poder sacar las probabilidades condicionales deseadas, para ello se selecciona tanto la variable de respuesta y la variable explicativa con el fin de saber que cuales son los que cumplen las dos condiciones, es decir, lo que tiene seguros de autos y lo que sean de empleo de gestión intermedia, pero que se cumplan al mismo tiempo, si por ejemplo tiene un seguro de autos y no tiene el empleo de gestión intermedia, en automático se descarta, hasta que se cumplan con las dos condiciones, las cuales son que tengan seguro y que tenga un empleo de gestión intermedio. Se hace así con cada una de las variables para poder sacar las probabilidades condicionales. “dependiendo el tipo de dato, el cálculo de  $P(X | Y)$  varia, existen diferentes variantes de Naive Bayes para manejar diferentes tipos de datos<sup>7</sup>” (Olguin A, 2024, p.1). Si bien esto es cierto, debemos de encontrar o tener en cuenta que existen varios cálculos para poder obtenerlos, por ejemplo, podemos encontrar al Naive Bayes Gaussiano y el multinomial.

Este método es el que se ocupa para sacar probabilidades condicionales, el que se menciona al inicio del artículo, una vez explicada un poco la teoría conforme al ejercicio que se está elaborando, veamos los resultados de  $P(\text{Seguro} | \text{Clase social, Educación, Trabajo})$ . La probabilidad de que el cliente haya comprado un seguro, dado que pertenece a una determinada clase social, determinado tipo de empleo y determinado nivel de empleo, pero para ello debemos sacarlos de manera individual y después multiplicar todas las probabilidades condicionales más significativas, para poder obtener una probabilidad condicional de más de una variable explicativa.

Primero iremos por el subconjunto de clase social.

1. MSKC: 0.514841
2. MSKB2: 0.509934
3. MSKA: 0.505877
4. MSKB1: 0.505258
5. MSKD: 0.496734

---

<sup>7</sup> Olguin, Alfredo. 2024. «Una explicación introductoria al teorema de Bayes». Diapositivas. Facultad de Economía UNAM. 5 de septiembre de 2024.



la variable que tiene más significancia del subconjunto de clases social es la de la clase c, por lo que podemos decir que la probabilidad de que el cliente haya comprado un seguro de auto, dado que pertenece a la clase social c es de 51.48%, después la clase B2, A, B1 y por último la D

Sigamos con el subconjunto de nivel de educación

1. MOPLLAAG: 0.513308
2. MOPLMIDD: 0.510465
3. MOPLHOOG: 0.508027

En este caso el nivel de educación que tiene más significancia es el de educación baja, por lo que la probabilidad de tener un seguro de auto dado que tenemos un nivel de educación baja es del 51.04%, después le sigue la educación media y educación alta

Por último, el subconjunto de tipo de empleo

1. MBERARBO: 0.511949
2. MBERMIDD: 0.508244
3. MBERARBG: 0.507626
4. MBERBOER: 0.493925
5. MBERZELF: 0.467595

La variable que tiene más significancia es la de los obreros no calificados, por lo que la probabilidad de que tenga un seguro de autos dado que tiene un trabajo de obrero no calificado es del 51.18%, después le sigue los de gestión intermedia, obreros calificados, agricultores, emprendedores.

En primera instancia podemos decir que las personas que son de clases bajas y niveles de educación baja, así como trabajos no calificados son las personas que tiene más probabilidad de obtener un seguro de autos, con otro enfoque, son personas que tal vez no tengan un buen ingreso y con un trabajo tal vez mal pagado por el nivel de educación y de trabajo, no se puede dar el lujo de tener un incidente sin tener al menos seguro que los respalde, es por ello que en un primer momento el público adecuado para hacer promociones en base a los niveles socioculturales de las personas que son de clase social C y dejar a un lado a los emprendedores

que son los que menos probabilidad tiene y que aparte son ellos lo que tiene un nivel de educación alta. Es por ello por lo que las personas objetivas son las de clase C, nivel educativo bajo y obreros no calificados.

Ahora seleccionado estas variables como las más significantes de los subconjuntos que hicimos para planificar las de una mejor manera y multiplicarlas para así obtener una probabilidad condicional con más de unas variables, quedaría de la siguiente manera.

Obrero no calificaron	Clase C	Educación Baja	Simultaneo
0.511949	0.514841	0.513308	0.135294

En la tabla anterior nos muestran las variables significativas y con el 13.53% es la probabilidad de que pasen los tres eventos simultáneamente, o sea que si cumple con esas condiciones habrá un margen bastante amplio para que esa persona tenga que comprar un seguro.

Ahora veamos la probabilidad complementaria, es decir,  $P(\text{NO Seguro} \mid \text{Clase social, Educación, Trabajo})$ .

Obrero no Calificado	Clase C	Educación baja	Simultaneo
0.488051	0.485159	0.486692	0.11524

En este caso se está haciendo las probabilidades condicionales, lo que está haciendo es la probabilidad de que no pase un evento, dado que se cumple varias condiciones, en este ejemplo es la probabilidad de que no se compren seguros dado que pertenece a la clase social C, que tiene un nivel de **educación** baja y que es obrero no calificado, esto no ayuda para determinar, de manera simplista, si nuestras probabilidades son buenas o en realidad son malas, aunque

existen otras más opciones de evaluar estas, esta es una de más muchas opciones. En este caso podemos ver que

- P (Seguro | Clase social, Educación, Trabajo): 0.135294
- P (No Seguro | Clase social, Educación, Trabajo): 0.11524

vemos que esta probabilidad de que pase es aún mayor que la probabilidad de que no pase, por lo que nuestro modelo es bueno en primera instancia.

### **Coeficientes:**

Para determinar cómo es que reaccionan las variables explicativas, ante la variable de respuesta se usó un método de regresión en el que se sacan los coeficientes obtenidos el siguiente resultado

- MBERARBO: -0.00072831
- MSKC: 0.00044636
- MOPLLAAG: -0.00073553

Podemos decir que, por cada unidad, en este caso, que todo se mantenga constante y aumente una unidad los niveles de obreros no calificados, lo seguros van a disminuir un 0.0007 unidades, aunque es mínimo, pero esto quiere decir que, si los obreros siguen llegando y todo siga igual, estos tendrán una pérdida, por lo que se pueden emplear estrategias necesarias para mitigar esto y que no se generen pérdidas. Por otro lado, en la clase social C, nos dice que si aumenta una unidad de estas personas que pertenecen a la clase social C, lo seguros tendrán un aumento de 0.0004. Por último, en el coeficiente de el nivel de educación baja es que por cada unidad de personas que carecen de una educación buena, cuando aumentan las personas con niveles de educación baja, lo seguros tendrán una pérdida de la misma magnitud de que de las personas que son obreros no calificados, es decir de 0.0007

Estos son muy importantes, ya que con estos podemos hacer estrategias de saber que tantas afectas las variables cuando todo lo demás permanece constante, es decir que, que tanta sensibilidad tienen los seguros antes las variables seleccionadas y ver si generen pérdidas o no.

### **Precisión:**

Para saber la presión del modelo se emplearon paqueterías en Python sobre aprendizaje automatizado, más en específico la paquetería de Scikit-learn, para poder el resultado de la precisión del modelo es del 0.5002862049227247 lo que quiere decir que el modelo tiene un 50.03% de poder decir cual el valor en un futuro, si bien esto tiene un sesgo y existen aún más parámetros para poder determinar la efectividad del modelo, podemos decir que el modelo está bien de momento. También se hizo un reporte de clasificación como el siguiente:

Reporte de clasificación:

	precisión	recall	f1-score	support
0	0.51	0.71	0.59	868
1	0.49	0.32	0.39	801

En este reporte podemos ver que el 51% fueron a la clase 0, es decir a las personas que no tuvieron un seguro o que simplemente no pertenecían a las demás variables explicativas, por otro lado, vemos que en el 1 que son las que al menos pertenecen a una de las variables explicativas. En la segunda columna podemos apreciar que el 71% fueron las personas que fueron identificadas correctamente, y el 32% que fueron clasificadas en el 1, para la columna de f1-score, podemos observar que el 0 tiene un rendimiento moderado y que el de la clase 1 tiene un rendimiento muy bajo, la última columna de support es el número de muestras reales de cada clase en el conjunto de pruebas, en el caso del 0 fue la que se detuvo a tener más muestras y que la de 1 fue la que tuvo un desarrollo muy bueno.

En resumen, podemos ver que la clase 0 fue la que tuvo un buen desempeño ante las demás clases, en este caso, ante la clase número 1, dado que la clase 0 fue la que tuvo mayor precisión en el modelo, además de que se identificó más esta clase y aparte de que tuvo un mejor rendimiento ante el 1.

Podemos decir que aunque este trabajo práctico tenía más variables y se determinaron solo algunas, que además de que se hicieron en subconjuntos, esto para poder determinar aún mejor las significantes y que con esto, pudimos hacer el teorema de Bayes, el cual consiste en que un evento suceda dado que se lleva a cabo una serie de filtros, es decir, que se cumpla la condición de que las personas tengan un seguro de autos dado que tiene un determinado tipo de empleo,

un nivel de educación determinado y que pertenece a una clase social, es por ello que con el teorema de bayes pudimos salir un poco del paradigma que se tiene con los métodos que usan las economías más arraigadas ya que estos son lo que suelen ocupar métodos frecuentistas que a veces no se logra analizar de una buena manera como es que se eligen las variables o como es que es tan preciso un modelo, por ejemplo, en este método fuimos seleccionando variable tras variable para ver las más significativas y que se ajustara de mejor manera al modelo, cosa que métodos como regresión lineal múltiple no ocuparía, ya que ellos meterían las variables una tras otra aunque no tengan relación alguna con el modelo, si bien todas las matemáticas, y la teoría económica tiene sesgos en algunas cosas esto no quiere decir que con ello podemos saltarnos de manera absurda la lógica de meter variables más significativas.

### **Resultado con la Red Neuronal:**

Como se menciona con anterioridad, también se usó una red neuronal para poder determinar que variables son significativas en nuestros conjuntos establecidos para nuestro modelo y determinar la compra de seguros de autos.

Podemos ver que la variable de nuestro conjunto de clase social es la Clase social C (MSKC), en el conjunto de trabajo es la variable de obreros cualificados, por último, en nuestra variable de estudios es el nivel de estudios bajos.

Nivel de Educación	Tipo de empleo	Clase social
MOPLLAAG: 0.015794	MBERARBG: 0.017339	MSKC: 0.021974
MOPLMIDD: 0.015193	MBERZELF: 0.015708	MSKB2 0.020944
MOPLHOOG: 0.008584	MBERARBO: 0.013562	MSKD 0.020687
	MBERMIDD: 0.009957	MSKA 0.018026
	MBERBOER: -0.001545	MSKB1 0.012961

Aunque este modelo no mide precisamente la probabilidad de conseguir un determinado seguro de autos dado una serie de condiciones específicas, esta red neuronal mide cuál es la importancia de que cada variable sea relevante para que se consiga un seguro de autos, por medio de una asignación de pesos en una capa oculta, nos sigue dando que la clase social que

más determina esta compra de seguros es la clase c, sin embargo, los empleos cambian ya que en un teorema de bayes nos dice que un obrero no cualificado es quien compra mas seguros, pero con una red neuronal nos dice que son los obreros cualificados son quienes compran más los seguros de autos, por último, la clase social baja sigue siendo relevante en una red neuronal y en un teorema de bayes.

### **Precisión de una red neurona.**

Para poder hacer una comparación de que tan bien es nuestro modelo, tanto en una red neuronal como en un teorema de bayes, debemos de saber que la precisión del modelo de red neuronal el cual es:

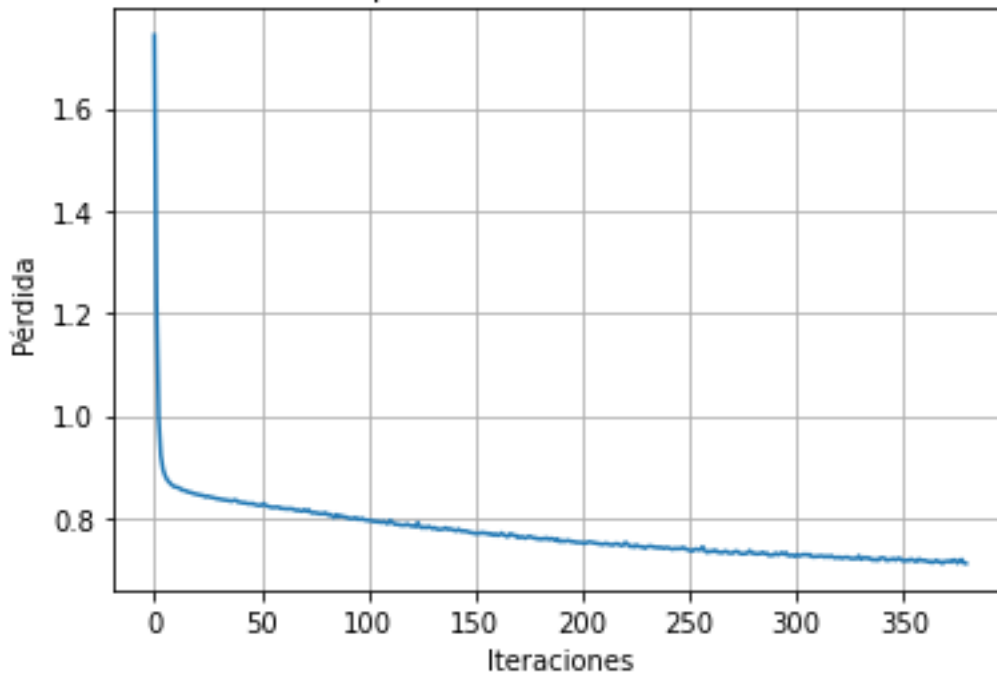
**Precisión: 0.5030042918454936**

Un poco menos a la hecha por el teorema de bayes que fue del 0.51, este modelo de una red neuronal no fue precisamente bueno para nuestro trabajo, esto demuestra que no siempre este tipo de modelo es bueno para un trabajo, a veces es mejor hacerlo de forma sencilla con otro tipo de modelos, como el teorema de bayes, que puede ser mejor para determinados trabajos.

### **Perdida de la información:**

Para poder hacer un modelo en una red neuronal es necesario entrenar la red con los datos de entrada, repitiendo el proceso con un determinado numero de iteraciones para que el modelo pueda ser mejor predictor, sin embargo, la red neurona puede llegar perder información en este proceso de aprendizaje, la red neuronal trata de minimizar está perdida de información, en este caso de trabajo respecto a los seguros de autos, la perdida de información fue la siguiente

Grafica 1: Evolución de la pérdida de información durante el entrenamiento



Se hicieron alrededor de 500 iteraciones para que el modelo pudiera aprender mejor, la pérdida de información o bien, la información que no se pudo capturar de una mejor, ya sea por problemas en la base de datos o de nuestro modelo mismo, fue de 0.7108. lo cual es relativamente alto, además de que el modelo durante 10 iteraciones no mejoro en absoluto, por ello, el modelo no fue útil para este caso.

## Conclusión

Nuestras probabilidades arrojaron que las variables que más se ajusta en la adquisición de seguros fueron obreros no calificados, personas con educación baja y clase social C, esto quiza nos pueda decir que la probabilidad de que una persona pueda tener un seguro dado las variable anteriores es del 13.52%, lo que se recomendaría al banco o agencia de seguros que su público debe de ser ese, aunque también determinamos que las personas obreras no calificadas y educación baja, causan perdidas si todo permanece constante por lo que haría que dar medidas para mitigar ello. Al emplear una red neuronal, nuestro modelo no mejoro el trabajo, ya que las variables siguen siendo las mismas, solo que el empleo cambia, estos son los obreros cualificados, sin embargo, nuestro modelo no tuvo una precisión alta, fue menor al teorema de bayes, esto demuestra que no siempre las redes neuronales son las mejores, estás

pueden llegar en presentar fallos y dar predictores equivocados, a veces es recomendable emplear formas sencillas para cierto casos, y en este trabajo es mejor usar el teoremas de bayes.

#### Bibliografía:

- Izaurieta, Fernando. s. f. «Redes Neuronales Artificiales». *Universidad de Concepción*.
- Olguin, Alfredo. 2024. «Una explicación introductoria al teorema de Bayes». Diapositivas. Facultad de Economía UNAM. 5 de septiembre de 2024.
- Valdez y Alfaro, Irene Patricia. 2018. «PROBABILIDAD Y ESTADÍSTICA». Diapositivas. DVB. junio de 2018. [http://www.dcb.unam.mx/profesores/irene/Notas/Tema\\_2-1\\_pantcompl.pdf](http://www.dcb.unam.mx/profesores/irene/Notas/Tema_2-1_pantcompl.pdf).
- 1.17. Neural Network Models (Supervised)». s. f. Scikit-Learn. [https://scikit-learn.org/1.5/modules/neural\\_networks\\_supervised.html](https://scikit-learn.org/1.5/modules/neural_networks_supervised.html).