

Explotación de características de secuencias para su uso en sistemas de recomendación

Máster Universitario en Investigación e Innovación en Inteligencia Computacional y Sistemas Interactivos

Trabajo Fin de Máster

Ricardo Sánchez-Guzmán Hitti



Índice

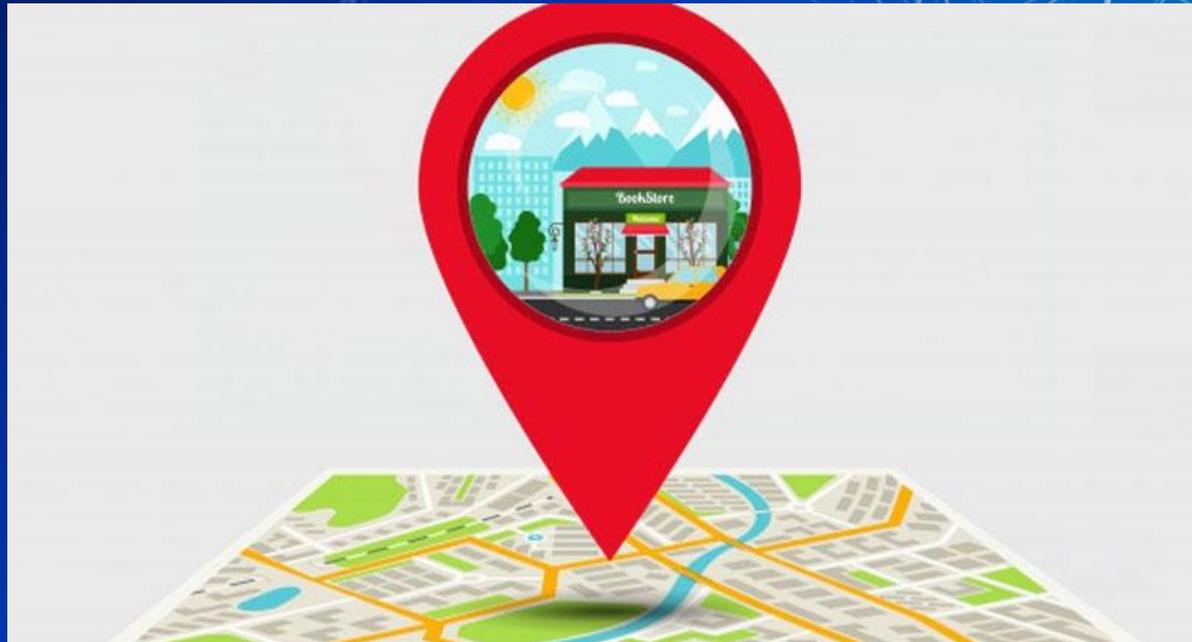
- Introducción
- Estado del arte
- Propuesta
- Diseño e implementación
- Experimentos y resultados
- Conclusiones y trabajo futuro

Índice

- **Introducción**
- Estado del arte
- Propuesta
- Diseño e implementación
- Experimentos y resultados
- Conclusiones y trabajo futuro

Introducción

Motivación

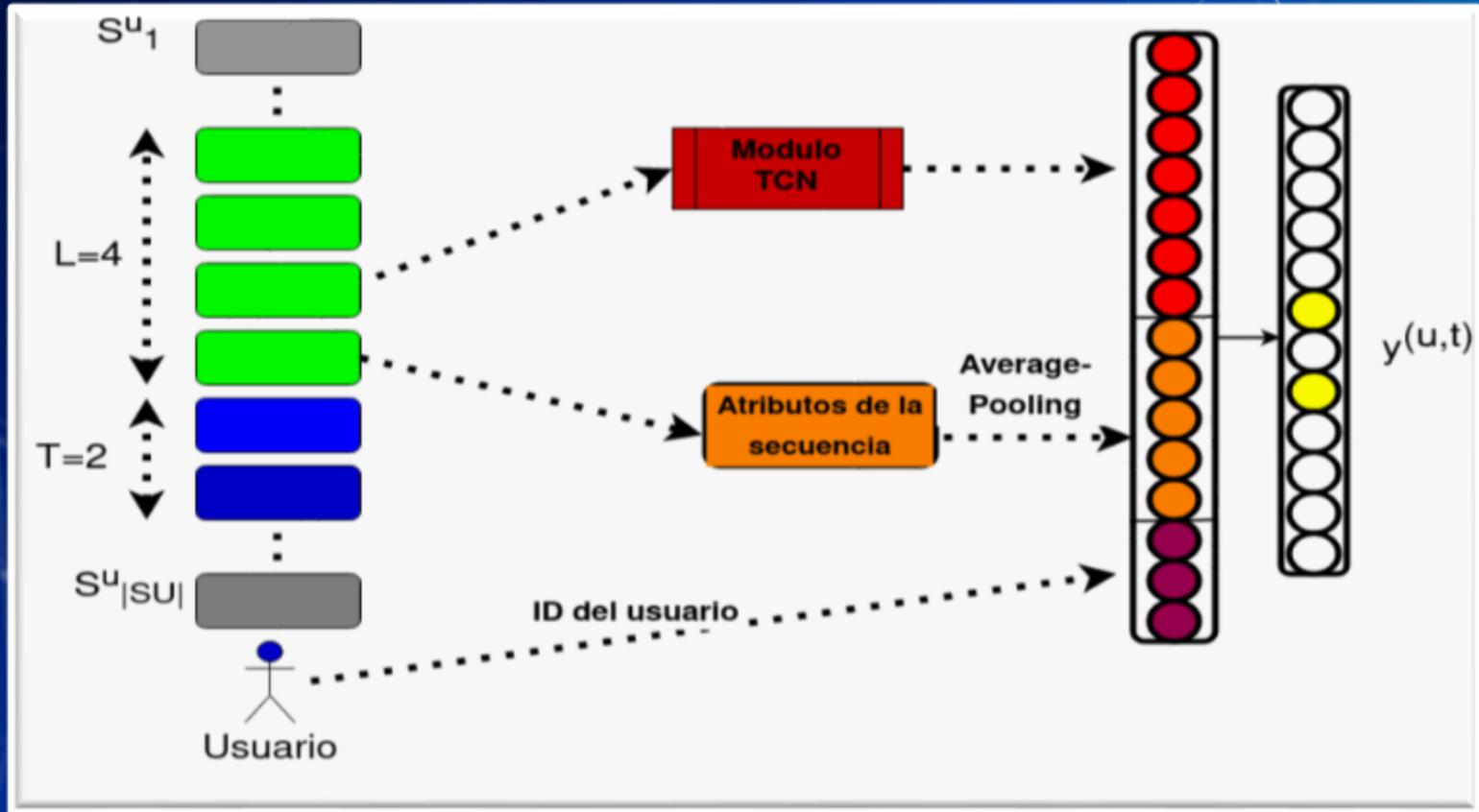


Introducción

Objetivos

- Plantear un nuevo modelo basado en redes neuronales que sea capaz de predecir el siguiente ítem en una secuencia.
- Analizar que influencia tienen cada uno de los atributos en el dominio turístico.
- Realizar una comparativa entre nuestro modelo y el resto de baselines del estado del arte.
- El uso de dos conjuntos de datos para evaluar el modelo.

Modelo propuesto



Índice

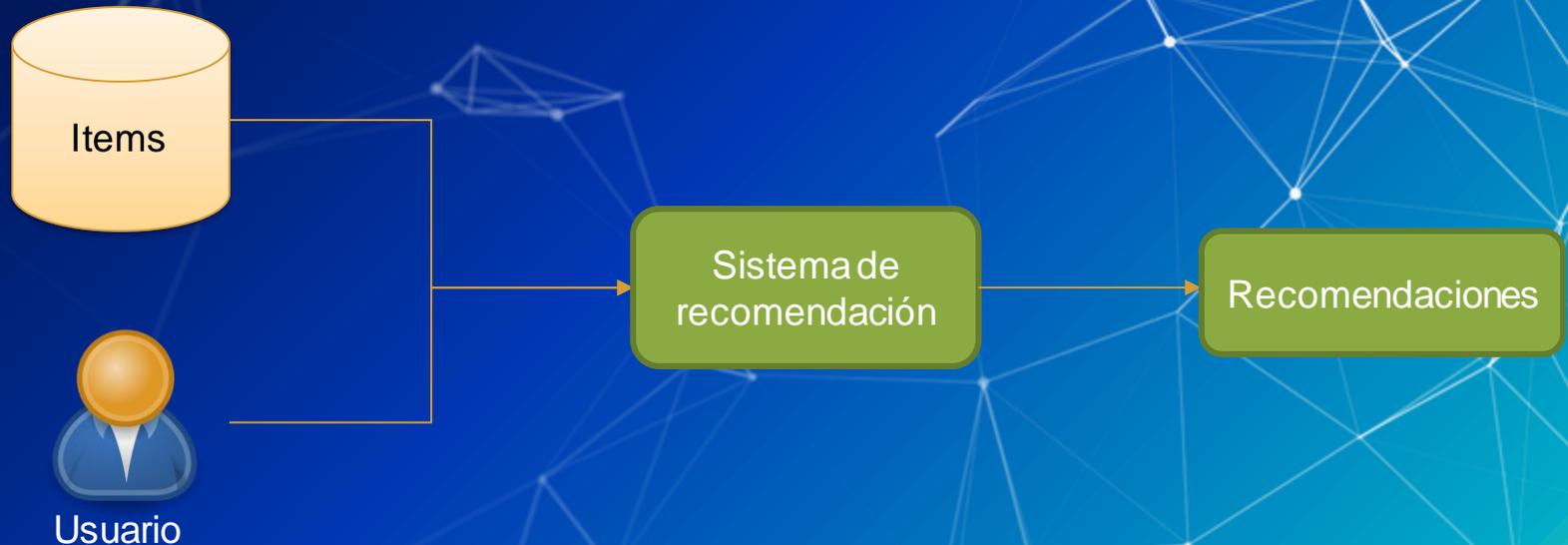
- Introducción
- **Estado del arte**
- Propuesta
- Diseño e implementación
- Experimentos y resultados
- Conclusiones y trabajo futuro

Sistemas de recomendación



Estado del arte

Sistema de recomendación



Estado del arte

Estrategias básicas de recomendación



Estado del arte

Fuentes de información alternativa

- Información temporal
- Información geográfica
- Información basada en comentarios

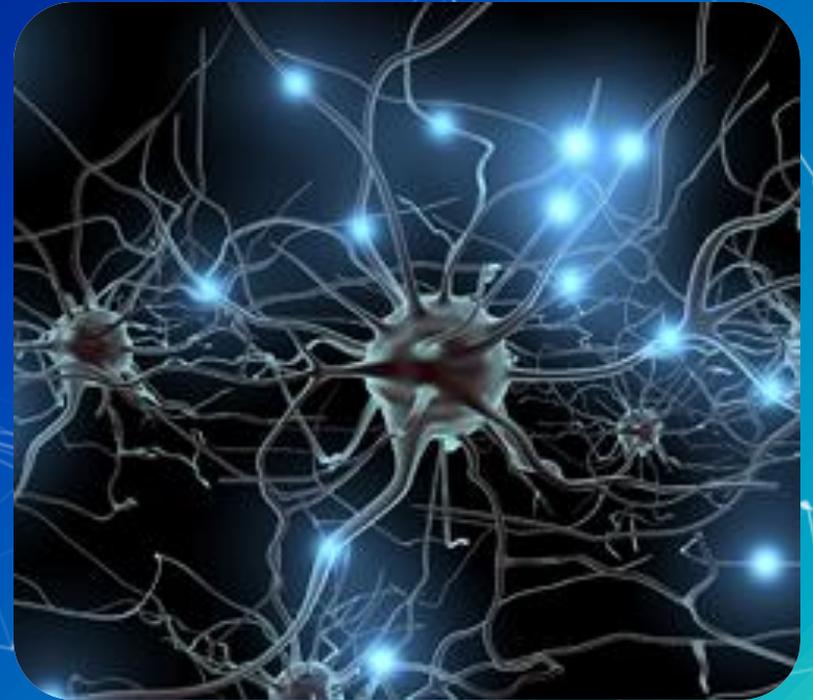


Dominio turístico

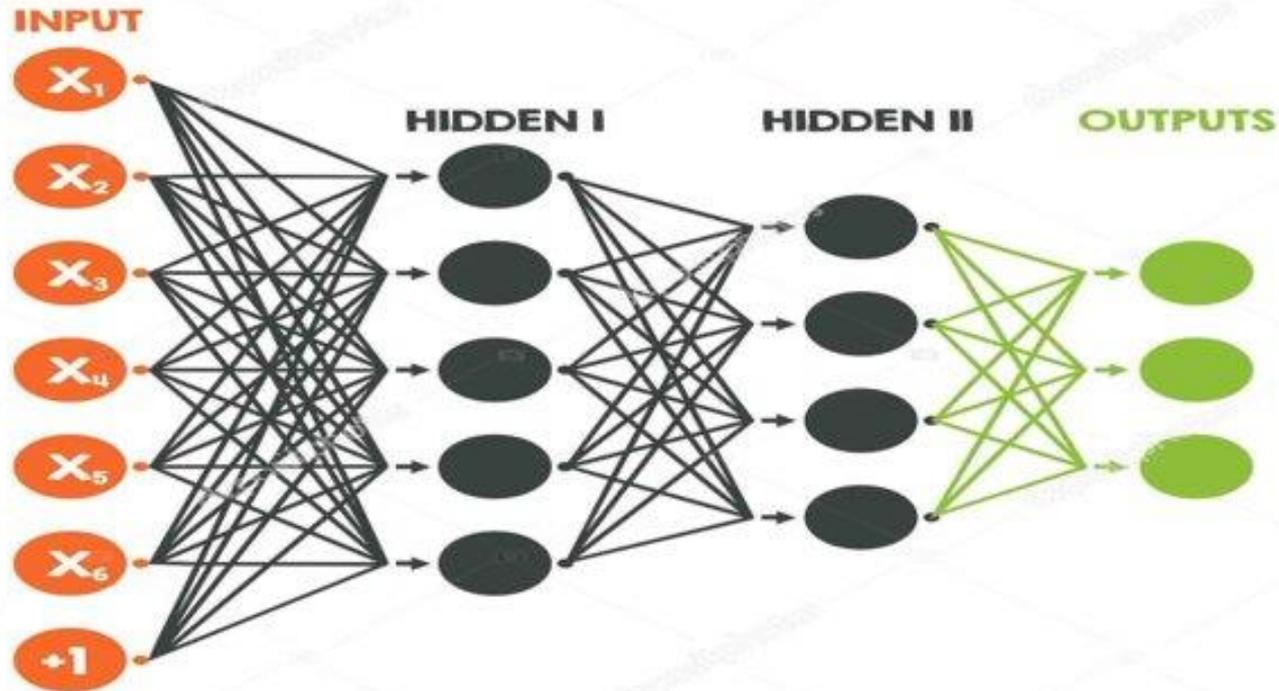


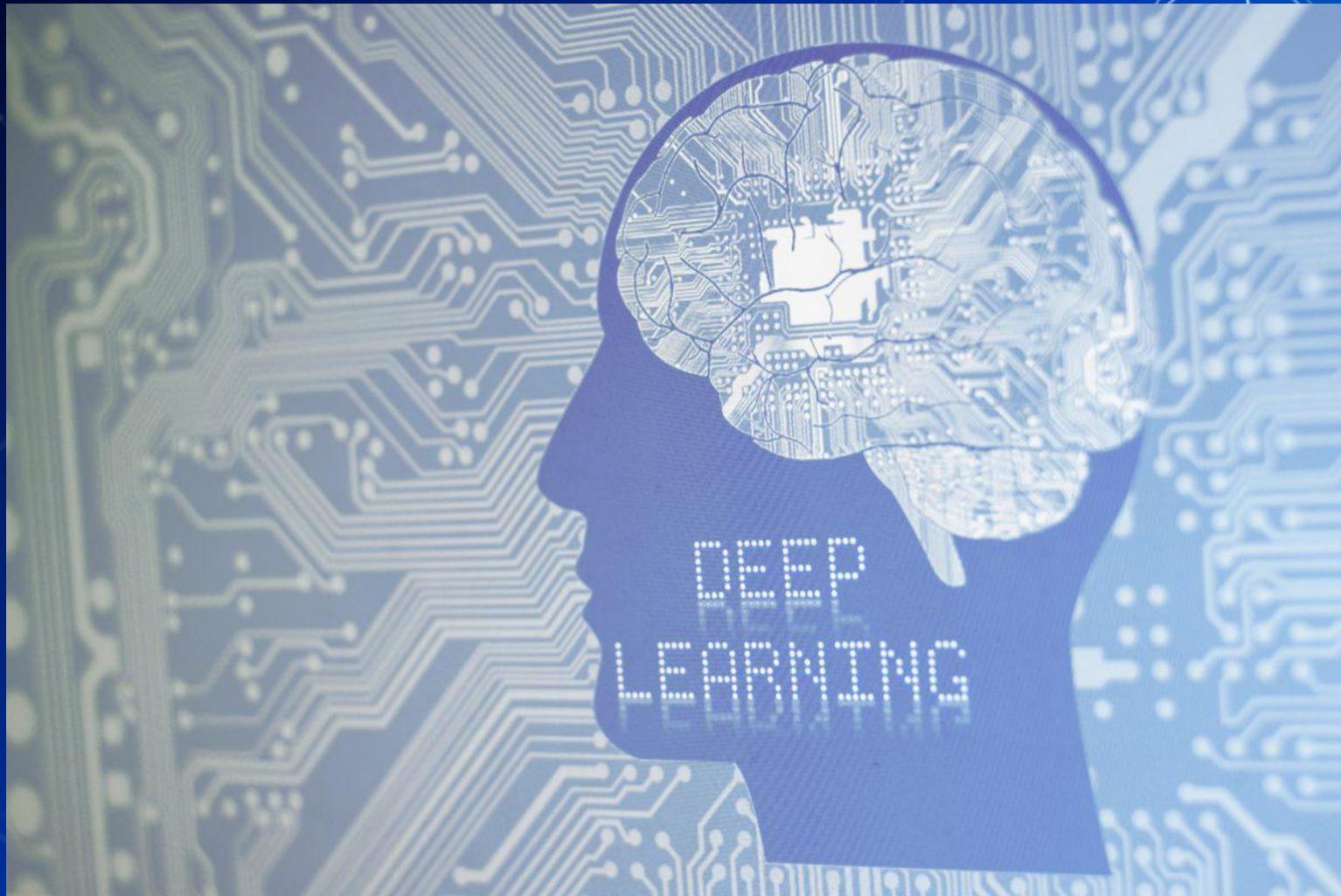


Redes neuronales



ARTIFICIAL NEURAL NETWORK

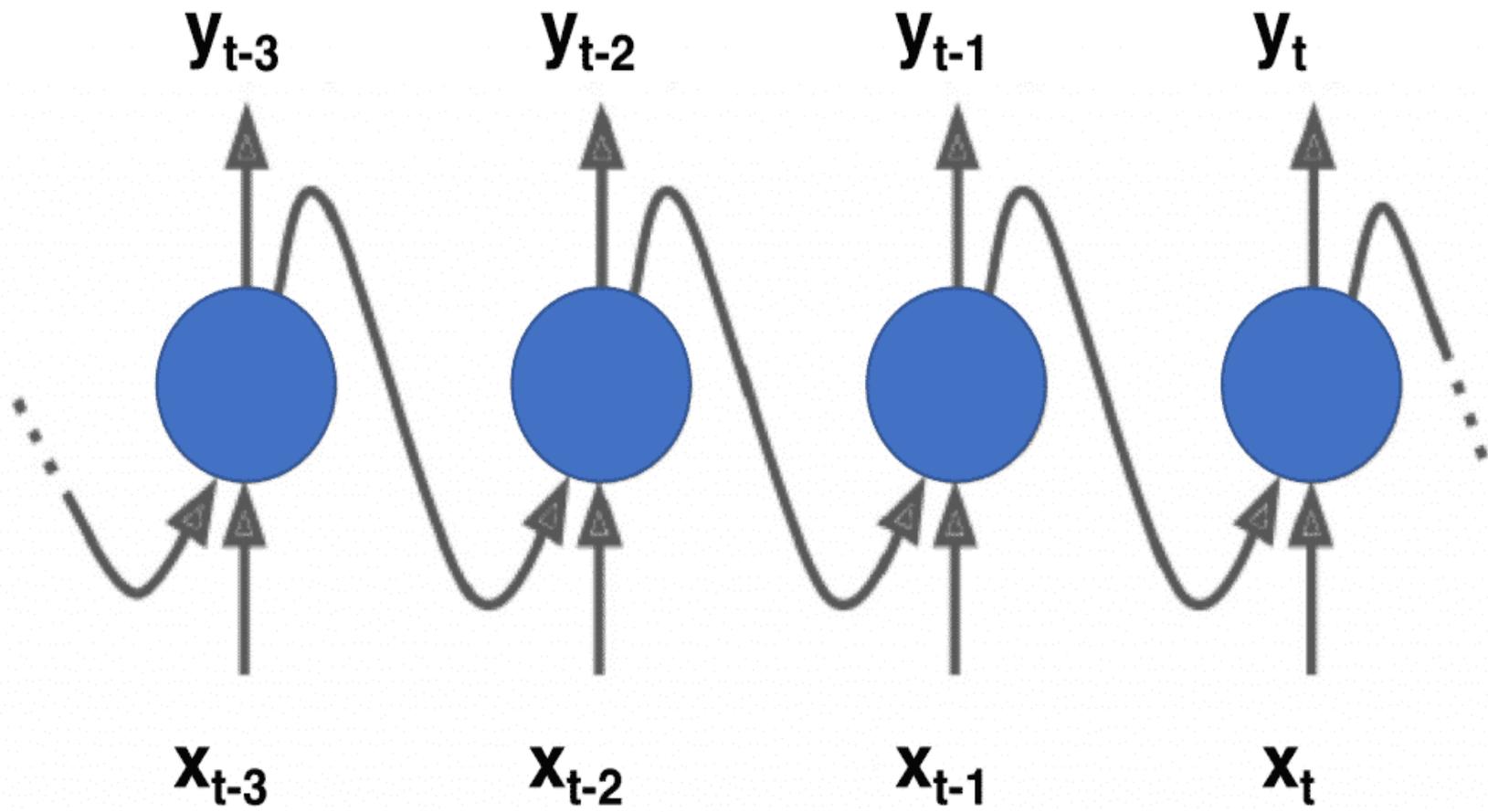




Estado del arte

Redes neuronales recurrentes (RNN)

- Concebidas en la década de 1980.
- La salida de una neurona se conecta de nuevo a la misma para mantener información de la salida anterior.
- No es capaz de soportar memoria a largo plazo.
- Suele aplicarse en listas y en datos temporales.



Estado del arte

Redes LSTM

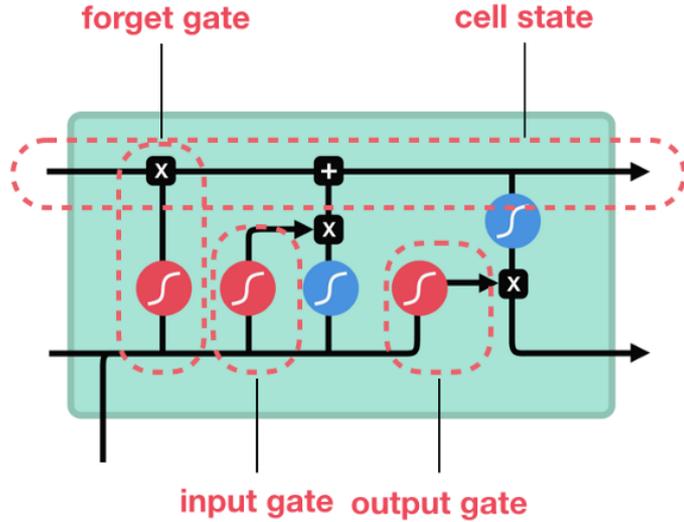
- Surgen en 1997.
- Extensión de las RNN para conseguir mantener información con grandes dependencias temporales.
- Cambia el concepto de neurona por celda de memoria.
- La celda de memoria contiene mecanismos para mantener, cambiar o eliminar el valor contenido en ella.
- Muy utilizadas en el reconocimiento de voz y la generación de textos.

Estado del arte

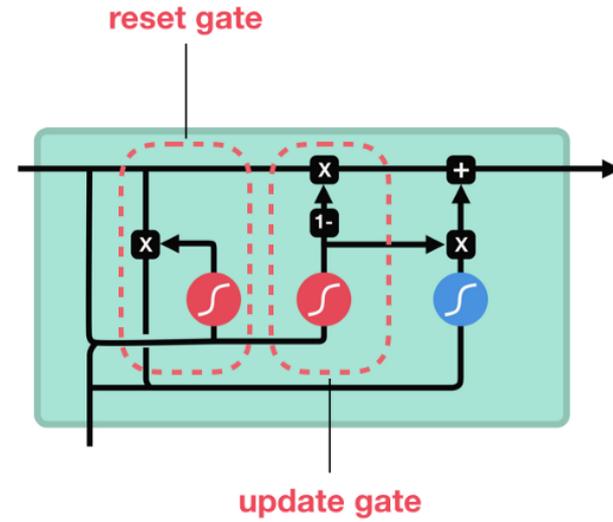
Redes GRU

- Aparecen en el año 2014.
- Similar a las redes LSTM ambas pueden capturar dependencias a largo plazo pero con menos parámetros a la hora de entrenar.

LSTM



GRU



sigmoid



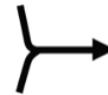
tanh



pointwise
multiplication



pointwise
addition

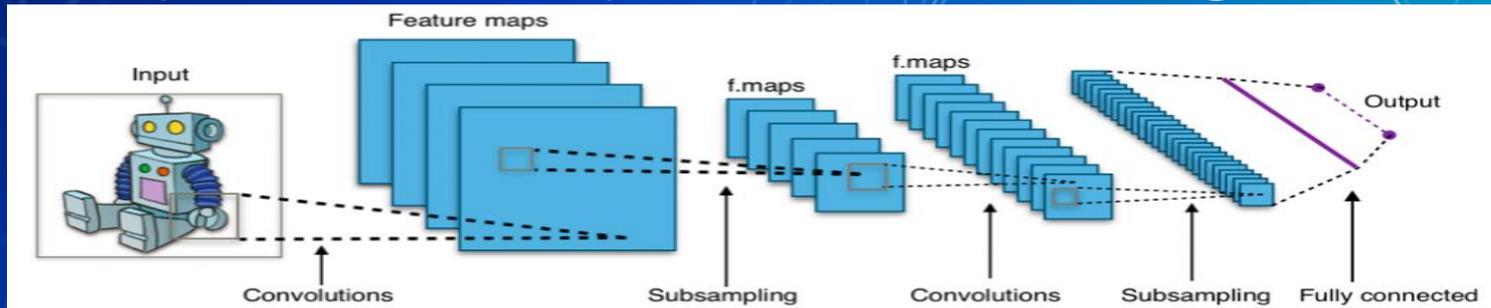


vector
concatenation

Estado del arte

Redes neuronales convolucionales (CNN)

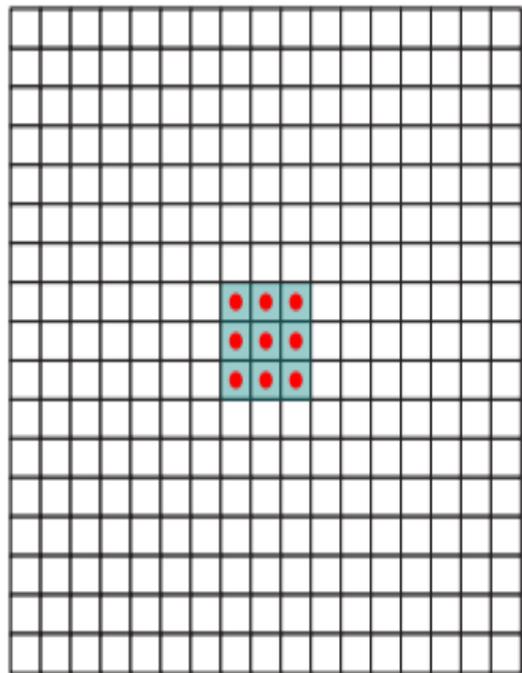
- Hacen su primera aparición en 1989.
- El nombre de la red viene dado por la operación de **convolución** denotada por $*$.
- Operaciones: convolución, pooling y aplanamiento.
- Muy utilizadas en el procesamiento de imágenes.



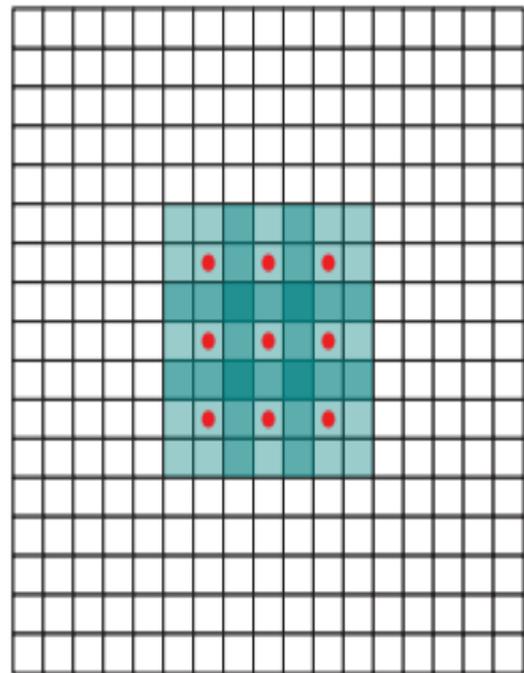
Estado del arte

Convolución dilatada

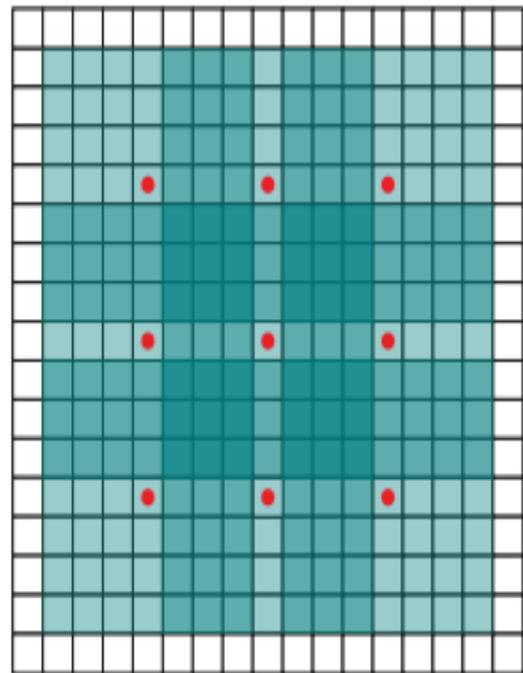
- La convolución dilatada se logra mediante el aumento del parámetro de D de dilatación del kernel.
- Gracias a esto expandimos el área del campo receptivo y permitimos involucrar mas información en la operación de convolución.
- El filtro crece, pero la cantidad de parámetros a ser ajustado no puesto que se rellenan con ceros los espacios.



(a)



(b)



(c)

Estado del arte

Capas de embeddings

- Mediante vectores compuestos de valores continuos podemos representar atributos.
- Muy utilizado en el ámbito del procesamiento del lenguaje natural.
- Menor coste que one-hot-encoder.

dog →

0.6	0.9	0.1	0.4	-0.7	-0.3	-0.2
-----	-----	-----	-----	------	------	------

Estado del arte

Word2Vec y fastText

Word2Vec (2013)

- Más rápido
- Existen redes pre-entradas con millones de palabras.
- Menor uso de memoria.



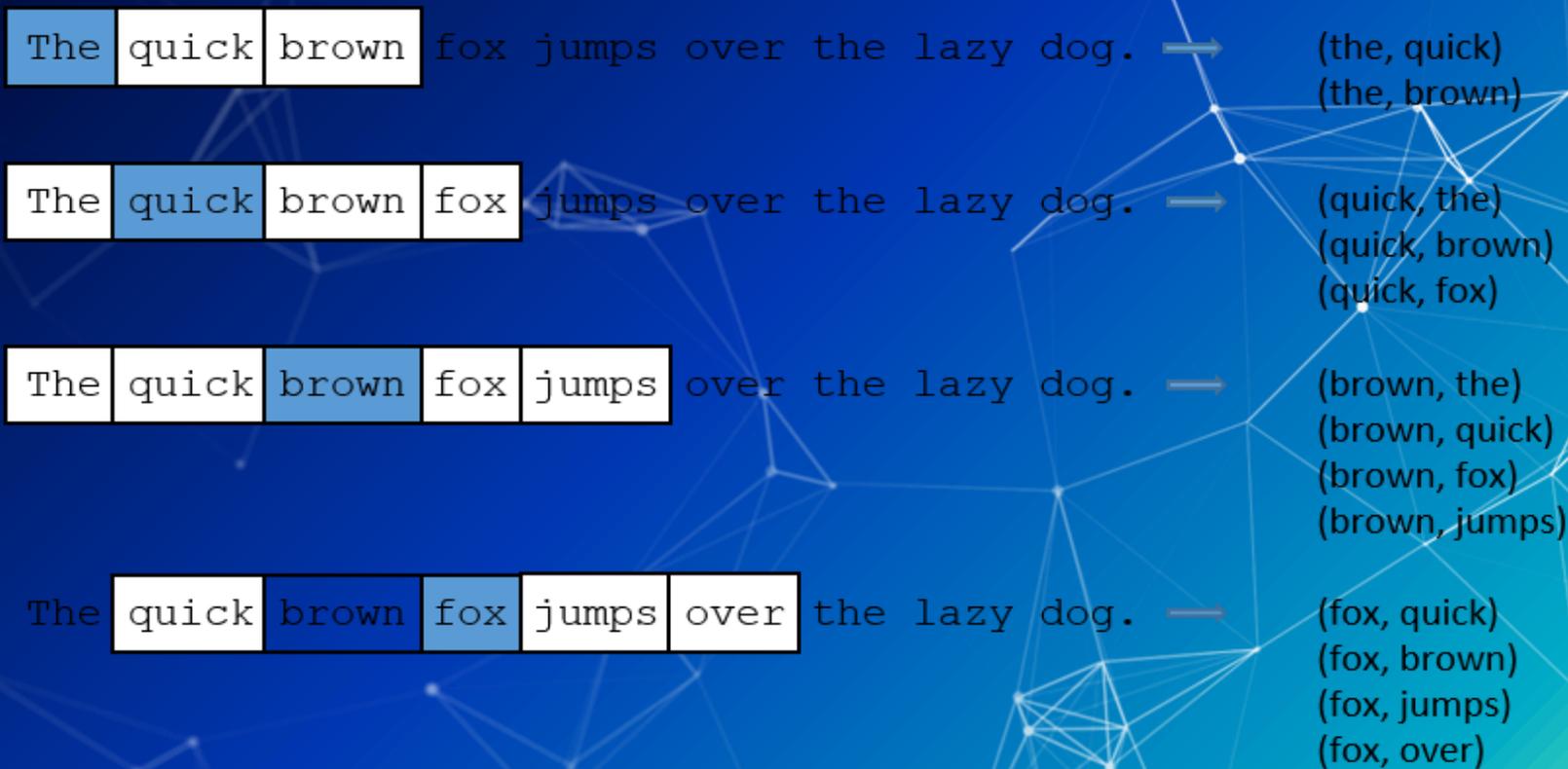
FastText (2015)

- Más robusto que Word2Vec ante palabras con erratas.
- Puede aprender palabras que no se encuentren en la fase de entrenamiento.

*fast*Text

Source Text

Training Samples



Sistemas de recomendación basados en secuencia

The image displays a collage of various recommendation systems. On the left, an Amazon-style interface shows 'Deals recommended for you' with product cards for a laptop, a smartwatch, and a tablet, each with a price and a 'Shop now' button. Below this are sections for 'Inspired by your shopping trends' (featuring books like 'Influence' and 'Flow') and 'Recommendations for you in Books' (featuring 'The Big Short' and 'Theory of Constraints'). In the center, a Netflix-style interface shows 'TV Shows' with a 'Genre' dropdown, 'Top Picks for Marcel' (featuring 'Lie to Me' and 'Star Trek: Voyager'), and 'Because you watched Dark' (featuring 'Black Mirror', 'The Frozen Dead', and 'Bates Motel'). Below these are 'New Releases' (featuring 'Orphan Black', 'Defenders', 'Frontier', 'Stranger Things', 'Train Fist', and 'Last Kingdom') and 'US TV Shows' (featuring 'Walking Dead', 'Madmen', 'The Mindy Project', 'Weeds', and 'Cosmos'). On the right, a dark-themed interface shows 'Playlists Made Just For You' with two playlist cards ('Dein Mix der Woche' and 'Dein Release Radar') and 'Top Recommendations For You' with a grid of movie and TV show thumbnails.

Estado del arte

Algoritmos



Estado del arte

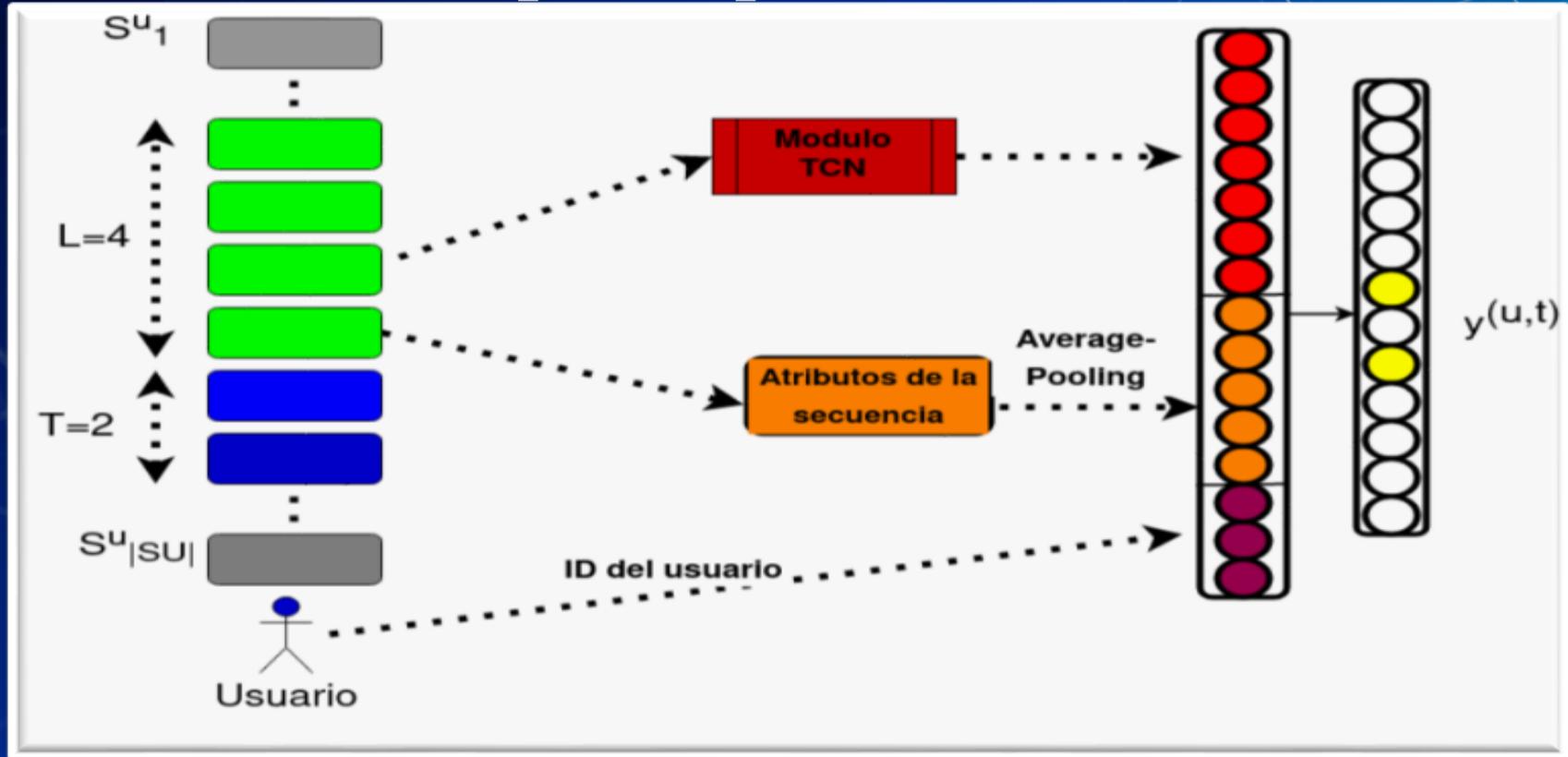
Conclusiones

- Hay muchos avances en las redes neuronales pero no todos se aplican a los sistemas de recomendación y más específicamente al dominio turístico.
- Es necesario tener en cuenta características como la fecha en la que se realiza un viaje o la zona que va a ir un usuario para poder mejorar las recomendaciones.

Índice

- Introducción
- Estado del arte
- **Propuesta**
- Diseño e implementación
- Experimentos y resultados
- Conclusiones y trabajo futuro

Modelo propuesto

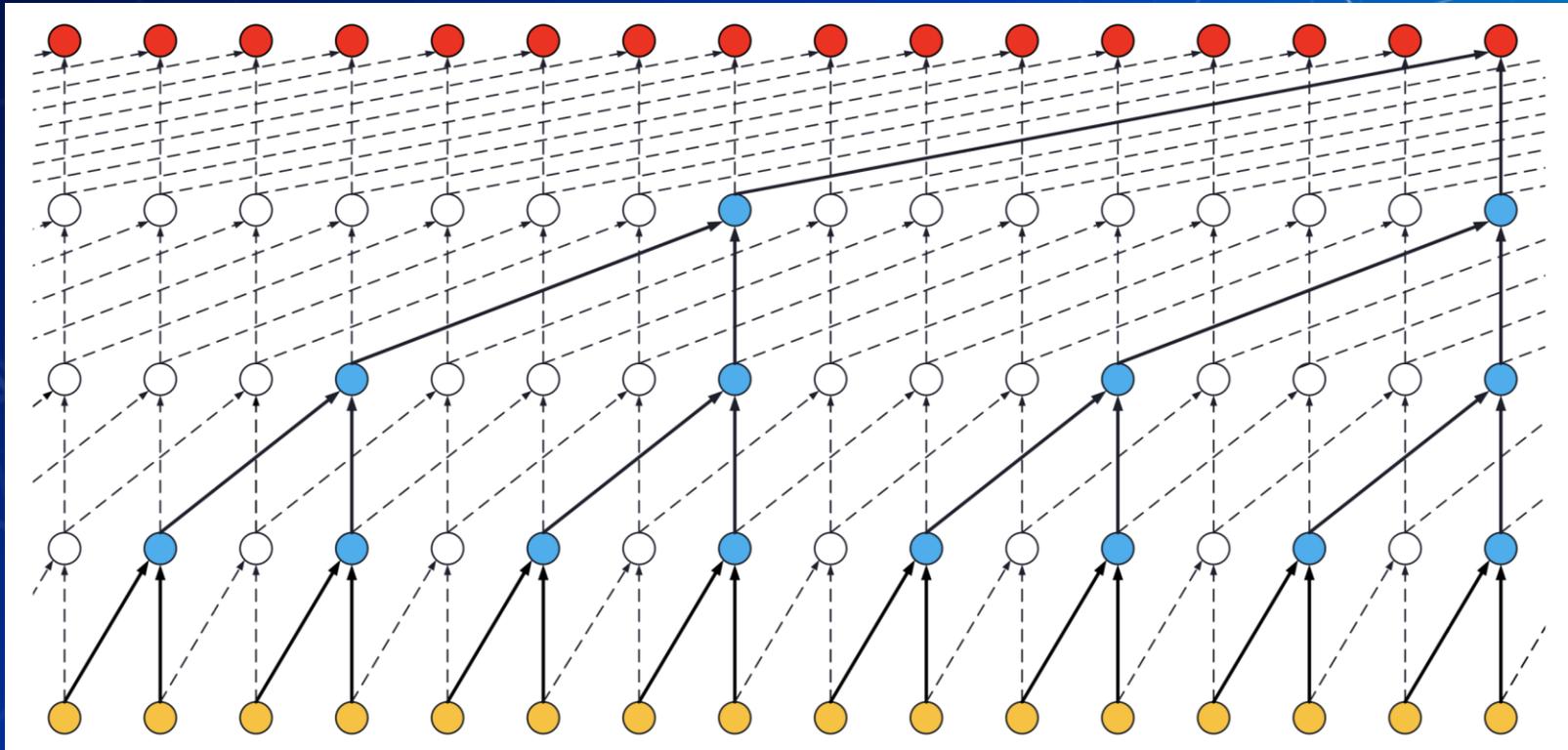


Descripción del Modelo

Características generales

- La secuencia de los POIs será la entrada del módulo TCN.
- Tratamiento de los atributos mediante las operaciones de máximo y media.
- Minimizar la función de pérdida con *Bayesian personalized ranking* (BPR).

Módulo TCN



Propiedades del módulo TCN

- Menor uso de memoria que las redes RNN.
- Flexibilidad a la hora de trabajar con el campo receptivo.
- Paralelismo.
- Capturar información tanto local como de grandes dependencias temporales.

Aumento de datos

- Elaborar subsecuencias a partir de una secuencia original
- Añadir ruido a la secuencia.



Original Session



Training Sequence 1

Training Sequence 2

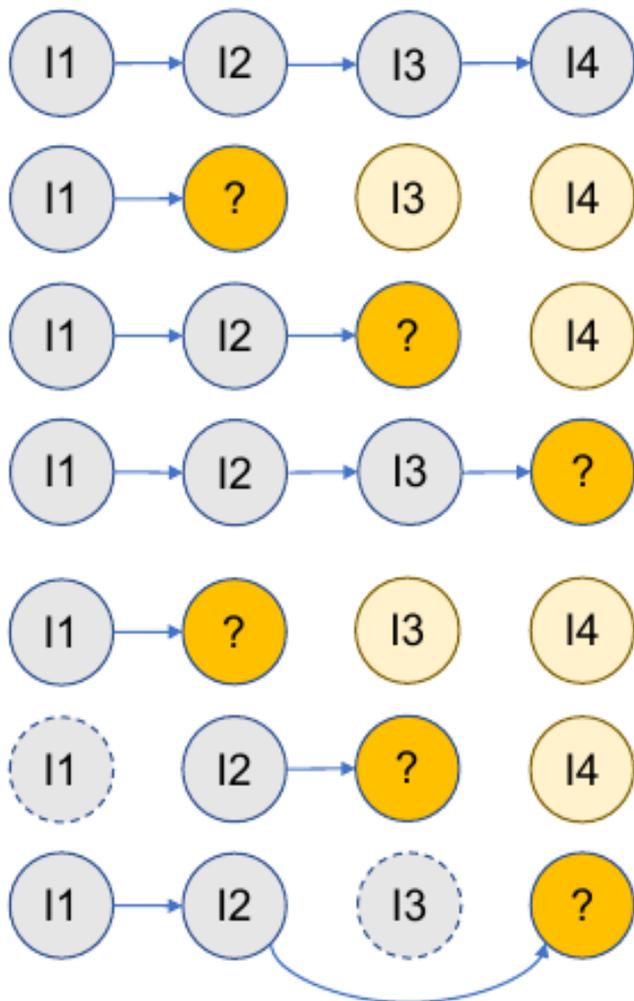
Training Sequence 3



Dropout Sequence 1

Dropout Sequence 2

Dropout Sequence 3



Tratamiento de características

- Realizar concatenación y multiplicación entre salidas de capas o vectores de embedding es muy costoso y en muchas ocasiones no garantiza buenos resultados.



Tratamiento de características

- Como alternativa a estas operaciones podemos capturar el maximo o la media de un conjunto de vectores de embedding.

0.5 | 0.1 | 0.5 | 0.9 | 0.1 | 0.4

0.3 | 0.2 | 0.4 | 0.6 | 0.3 | 0.7

0.4 | 0.1 | 0.6 | 0.3 | 0.8 | 0.2



0.5 | 0.2 | 0.6 | 0.9 | 0.8 | 0.7

0.5 | 0.1 | 0.5 | 0.9 | 0.1 | 0.4

0.3 | 0.2 | 0.4 | 0.6 | 0.3 | 0.7

0.4 | 0.1 | 0.6 | 0.3 | 0.8 | 0.2



0.4 | 0.1 | 0.5 | 0.6 | 0.4 | 0.4

Estrategias de entrenamiento

- **TOP1**

- La primera parte calcula la diferencia entre la muestra positiva y negativa y la segunda actúa como regularización.

- **BPR**

- Igual que en el caso anterior se utiliza la función sigmoideal sobre la diferencia entre la puntuación de los ejemplos negativos y positivos.

Descripción del Modelo

Características generales

- Memoria a largo plazo y personalización mediante el ID del usuario.
- Análisis de las secuencias mediante el módulo *Temporal Convolutional Networks*.
- Información auxiliar mediante embending de características que permite tener en cuenta atributos como la fecha o las coordenadas de un lugar.

Índice

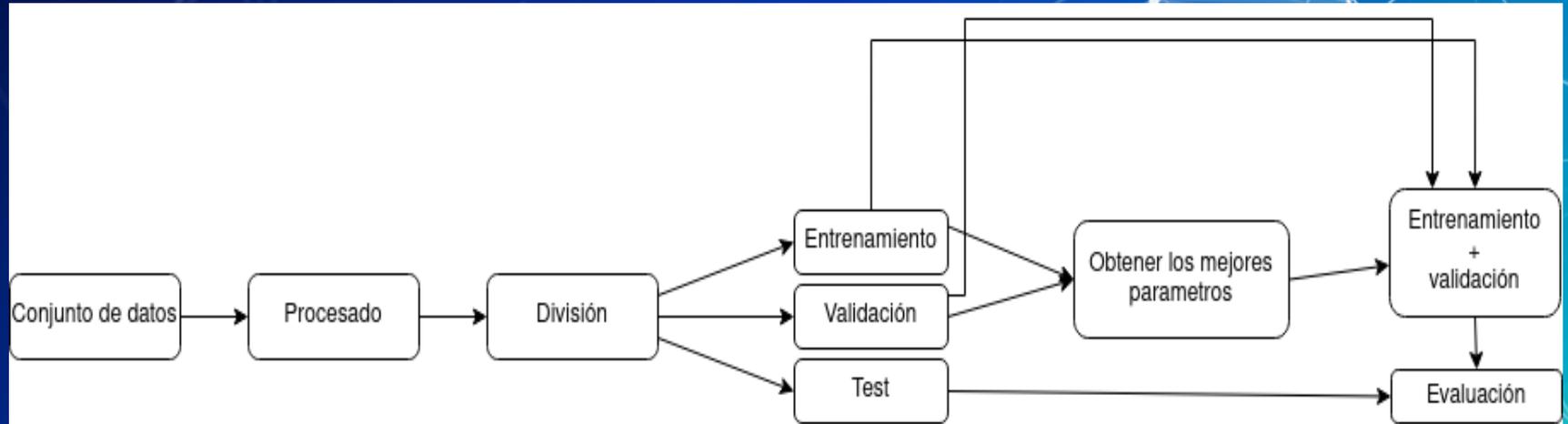
- Introducción
- Estado del arte
- Propuesta
- **Diseño e implementación**
- Experimentos y resultados
- Conclusiones y trabajo futuro

Librerías externas

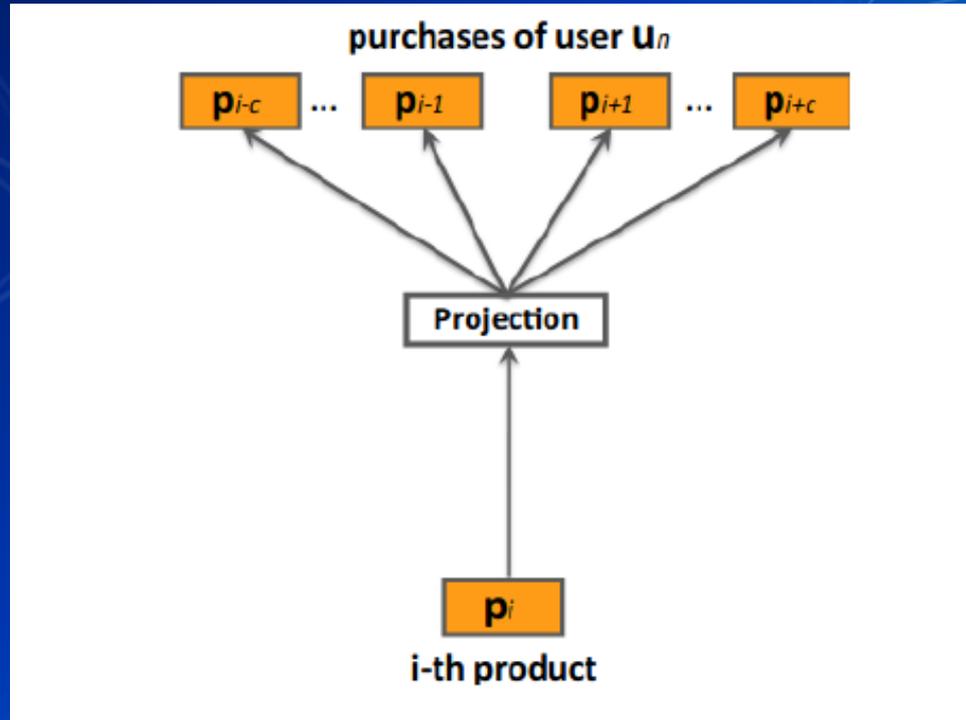
- Caser
 - Librería implementada en Pytorch.
- CaseRecommender
 - Librería utilizada para la ejecución de baselines básicos.
- Fossil
 - Librería utilizada para la ejecución de baselines secuenciales.
- TCN
 - Librería que implementa las redes convolucionales dilatadas en forma de bloque.

Metodología

Flujo de trabajo



Vectores pre-entrados Item2Vec



Resultados con Word2Vec y fastText

Conjuntos de datos	Modelos	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Foursquare	Word2Vec	0.0129	0.0111	0.0096	0.0129	0.0554	0.0759
	fastText	0.0264	0.0130	0.0088	0.0264	0.0549	0.0761
Yelp	Word2Vec	0.0129	0.0111	0.0096	0.0129	0.0554	0.0949
	fastText	0.0128	0.0105	0.0093	0.0128	0.0524	0.0921

Resultados con Word2Vec y fastText

Conjuntos de datos	Modelos	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Foursquare	Word2Vec	0.0129	0.0111	0.0096	0.0129	0.0554	0.0759
	fastText	0.0264	0.0130	0.0088	0.0264	0.0549	0.0761
Yelp	Word2Vec	0.0129	0.0111	0.0096	0.0129	0.0554	0.0949
	fastText	0.0129	0.0105	0.0096	0.0129	0.0524	0.0921

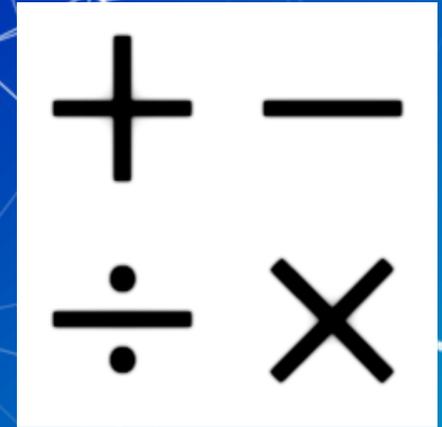
Índice

- Introducción
- Estado del arte
- Propuesta
- Diseño e implementación
- **Experimentos y resultados**
- Conclusiones y trabajo futuro

Pruebas

Descripción de las pruebas

- Análisis de dos conjuntos de datos, uno con escasa información como **Foursquare** y otro con información tanto geográfica como de opiniones de usuarios como **Yelp**.
- Comparativa con los baselines y nuestro modelo.
- Influencia de cada atributo en el modelo neuronal.
- Análisis de los hiperparámetros.
- Influencia de las operaciones máximo y media.
- Influencia de los atributos y del módulo TCN.



Estado del arte

Métricas para la evaluación de sistemas

- Recall

$$\text{Recall} = \frac{\#TP}{\#TP + \#FN}$$

- Precisión

$$\text{Precision} = \frac{\#TP}{\#TP + \#FP}$$



FOURSQUARE

Foursquare

Resultados obtenidos con los baselines

Tipo	Recomendador	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Básico	POP	0.0885	0.0428	0.0304	0.0231	0.0570	0.0789
	BPR	0.0200	0.0149	0.0123	0.0058	0.0200	0.0339
	ItemKNN	0.0163	0.0112	0.0094	0.0059	0.0192	0.0313
Secuencial	MC	0.0522	0.0440	0.0347	0.0092	0.0442	0.0703
	FPMC	0.0492	0.0434	0.0361	0.0116	0.0477	0.0790
	FOSSIL	0.0502	0.0442	0.0324	0.0374	0.0428	0.0796
Redes neuronales	Caser	0.0637	0.0390	0.0281	0.0177	0.0516	0.0721
	Cosrec	0.0652	0.0412	0.0274	0.0175	0.0554	0.0714
	GRU4Rec	0.0402	0.0160	0.0100	0.0402	0.0801	0.1006
	Nuestro modelo	0.0889	0.0495	0.0372	0.0234	0.0648	0.0846

Tabla 4.5: Resultados de los baselines y de nuestro modelo con el conjunto de datos de Foursquare.

Foursquare

Influencia de los atributos

Obtenemos los mejores resultados con el Tipo 3

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Atributos	Tipo 1	0.0830	0.0458	0.0312	0.0215	0.0598	0.0779
	Tipo 2	0.0836	0.0472	0.0320	0.0212	0.0626	0.0822
	Tipo 3	0.0859	0.0478	0.0320	0.0221	0.0628	0.0816
	Fecha	0.0851	0.0452	0.0315	0.0219	0.0588	0.0799
	Intervalo temporal	0.0843	0.0433	0.0304	0.0227	0.0554	0.0766
	Distancia geográfica	0.0757	0.0404	0.0287	0.0202	0.0523	0.0718
	Sin Módulo TCN	0.0739	0.0441	0.0309	0.0197	0.0587	0.0788
Modelo	Sin Atributos	0.0759	0.0456	0.0316	0.0202	0.0603	0.0804
	Mejor modelo (MAX)	0.0808	0.0425	0.0293	0.0218	0.0548	0.0727
	Mejor modelo (AVG)	0.0889	0.0495	0.0372	0.0234	0.0648	0.0846

Foursquare

Influencia de los atributos

El intervalo temporal supera a la distancia entre POIs.

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Atributos	Tipo 1	0.0830	0.0458	0.0312	0.0215	0.0598	0.0779
	Tipo 2	0.0836	0.0472	0.0320	0.0212	0.0626	0.0822
	Tipo 3	0.0859	0.0478	0.0320	0.0221	0.0628	0.0816
	Fecha	0.0851	0.0452	0.0315	0.0219	0.0588	0.0799
	Intervalo temporal	0.0843	0.0433	0.0304	0.0227	0.0554	0.0766
	Distancia geográfica	0.0757	0.0404	0.0287	0.0202	0.0523	0.0718
Modelo	Sin Módulo TCN	0.0739	0.0441	0.0309	0.0197	0.0587	0.0788
	Sin Atributos	0.0759	0.0456	0.0316	0.0202	0.0603	0.0804
	Mejor modelo (MAX)	0.0808	0.0425	0.0293	0.0218	0.0548	0.0727
	Mejor modelo (AVG)	0.0889	0.0495	0.0372	0.0234	0.0648	0.0846

Foursquare

Influencia de los atributos

El módulo TCN parece ser más relevante en el modelo que los atributos

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Atributos	Tipo 1	0.0830	0.0458	0.0312	0.0215	0.0598	0.0779
	Tipo 2	0.0836	0.0472	0.0320	0.0212	0.0626	0.0822
	Tipo 3	0.0859	0.0478	0.0320	0.0221	0.0628	0.0816
	Fecha	0.0851	0.0452	0.0315	0.0219	0.0588	0.0799
	Intervalo temporal	0.0843	0.0433	0.0304	0.0227	0.0554	0.0766
	Distancia geográfica	0.0757	0.0404	0.0287	0.0202	0.0523	0.0718
Modelo	Sin Módulo TCN	0.0739	0.0441	0.0309	0.0197	0.0587	0.0788
	Sin Atributos	0.0759	0.0456	0.0316	0.0202	0.0603	0.0804
	Mejor modelo (MAX)	0.0808	0.0425	0.0293	0.0218	0.0548	0.0727
	Mejor modelo (AVG)	0.0889	0.0495	0.0372	0.0234	0.0648	0.0846

Foursquare

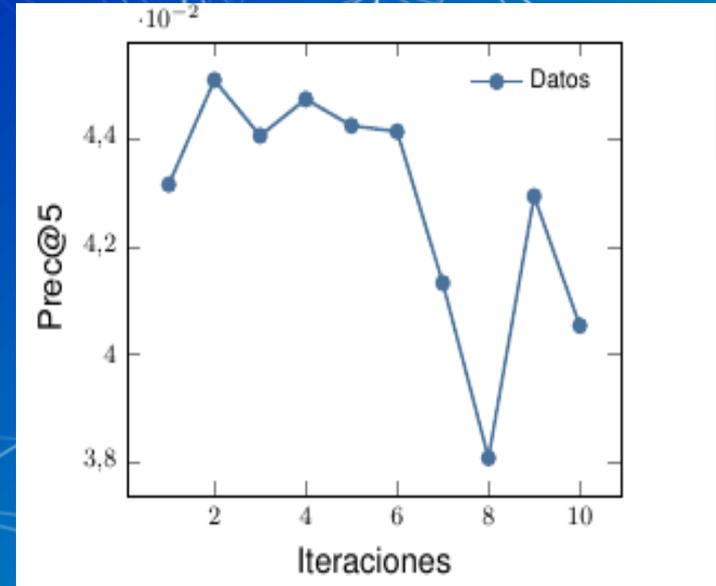
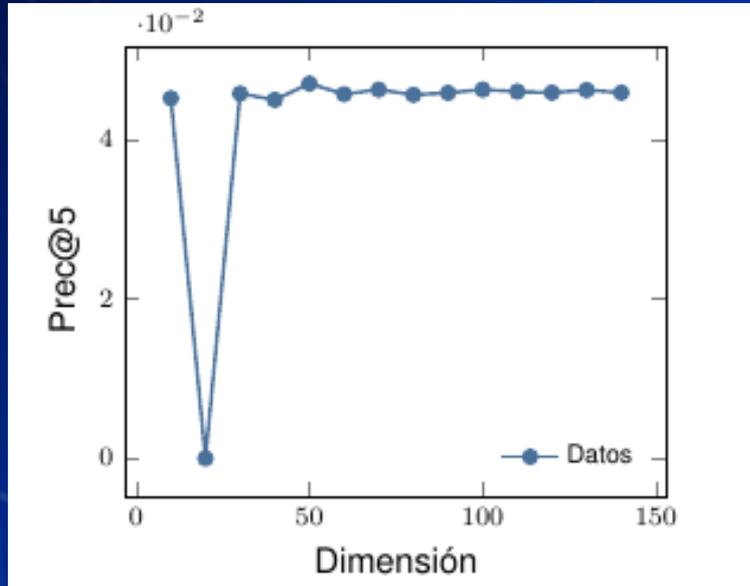
Influencia de los atributos

La operación media sobre los vectores de embedding obtiene los mejores resultados

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Atributos	Tipo 1	0.0830	0.0458	0.0312	0.0215	0.0598	0.0779
	Tipo 2	0.0836	0.0472	0.0320	0.0212	0.0626	0.0822
	Tipo 3	0.0859	0.0478	0.0320	0.0221	0.0628	0.0816
	Fecha	0.0851	0.0452	0.0315	0.0219	0.0588	0.0799
	Intervalo temporal	0.0843	0.0433	0.0304	0.0227	0.0554	0.0766
	Distancia geográfica	0.0757	0.0404	0.0287	0.0202	0.0523	0.0718
Modelo	Sin Módulo TCN	0.0739	0.0441	0.0309	0.0197	0.0587	0.0788
	Sin Atributos	0.0759	0.0456	0.0316	0.0202	0.0603	0.0804
	Mejor modelo (MAX)	0.0808	0.0425	0.0293	0.0218	0.0548	0.0727
	Mejor modelo (AVG)	0.0889	0.0495	0.0372	0.0234	0.0648	0.0846

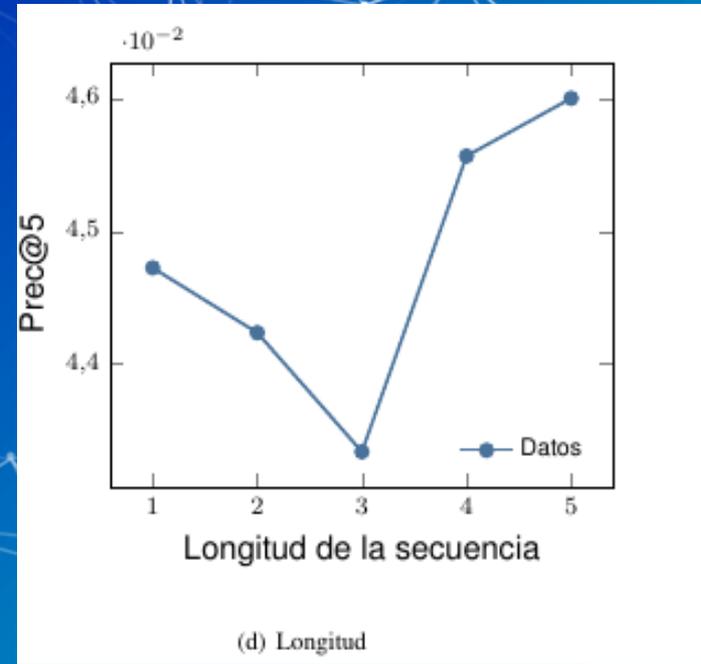
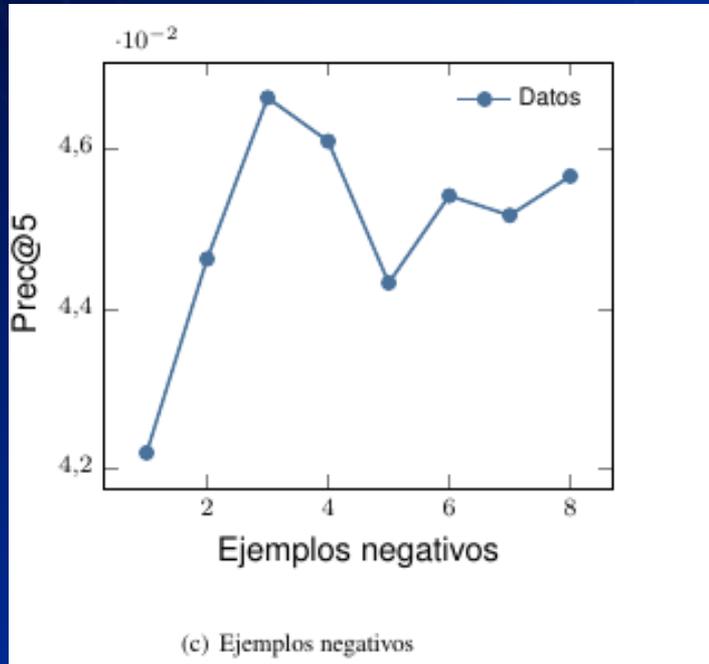
Foursquare

Análisis de hiperparámetros



Foursquare

Análisis de hiperparametros





Yelp

Resultados obtenidos con los baselines

Tipo	Recomendador	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Básico	POP	0.0044	0.0049	0.0046	0.0020	0.0113	0.0207
	BPR	0.0225	0.0185	0.0164	0.0100	0.0418	0.0737
	ItemKNN	0.0315	0.0247	0.0204	0.0149	0.0541	0.0940
Secuencial	MC	0.0256	0.0204	0.0180	0.0098	0.0396	0.0691
	FPMC	0.0306	0.0256	0.0227	0.0120	0.0501	0.0874
	FOSSIL	0.0294	0.0242	0.0211	0.0110	0.0466	0.0812
Redes neuronales	Caser	0.0311	0.0260	0.0235	0.0119	0.0491	0.0874
	Cosrec	0.0316	0.0273	0.0239	0.0117	0.0504	0.0880
	GRU4Rec	0.0113	0.0087	0.0075	0.0113	0.0435	0.0758
	Nuestro modelo	0.0386	0.0315	0.0275	0.0149	0.0597	0.1014

Tabla 4.7: Resultados de los baselines y de nuestro modelo con el conjunto de datos de Yelp. En negrita, el mejor resultado para cada métrica.

Yelp

Influencia de los atributos

La fecha de cuando se realizó el comentario es el atributo más relevante dentro del bloque de comentarios

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Comentarios	Utilidad	0.0328	0.0264	0.0231	0.0129	0.0508	0.0877
	Gracioso	0.0325	0.0264	0.0232	0.0125	0.0502	0.0875
	Interesante	0.0316	0.0269	0.0237	0.0121	0.0511	0.0890
	Fecha	0.0332	0.0271	0.0240	0.0128	0.0511	0.0911
	<i>Fusión</i>	0.0348	0.0295	0.0259	0.0132	0.0551	0.0961
Geografía	Dirección	0.0339	0.0276	0.0250	0.0134	0.0528	0.0938
	Ciudad	0.0343	0.0283	0.0249	0.0129	0.0535	0.0939
	Estado	0.0336	0.0270	0.0239	0.0129	0.0513	0.0903
	Código postal	0.0358	0.0294	0.0255	0.0137	0.0565	0.0952
	<i>Fusión</i>	0.0360	0.0290	0.0255	0.0139	0.0552	0.0981
Intervalos	Intervalo temporal	0.0304	0.0251	0.0221	0.0114	0.0473	0.0824
	Distancia geográfica	0.0287	0.0238	0.0213	0.0107	0.0447	0.0795
Modelo	Sin Módulo TCN	0.0345	0.0265	0.0231	0.0132	0.0500	0.0867
	Sin Atributos	0.0335	0.0283	0.0249	0.0128	0.0536	0.0933
	Mejor modelo (MAX)	0.0340	0.0283	0.0251	0.0132	0.0542	0.0942
	Mejor modelo (AVG)	0.0386	0.0315	0.0275	0.0149	0.0597	0.1014

Yelp

Influencia de los atributos

El código postal del lugar de interés es el más relevante dentro del bloque de geografía

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Comentarios	Utilidad	0.0328	0.0264	0.0231	0.0129	0.0508	0.0877
	Gracioso	0.0325	0.0264	0.0232	0.0125	0.0502	0.0875
	Interesante	0.0316	0.0269	0.0237	0.0121	0.0511	0.0890
	Fecha	0.0332	0.0271	0.0240	0.0128	0.0511	0.0911
	<i>Fusión</i>	0.0348	0.0295	0.0259	0.0132	0.0551	0.0961
Geografía	Dirección	0.0339	0.0276	0.0250	0.0134	0.0528	0.0938
	Ciudad	0.0343	0.0283	0.0249	0.0129	0.0535	0.0939
	Estado	0.0336	0.0270	0.0239	0.0129	0.0513	0.0903
	Código postal	0.0358	0.0294	0.0255	0.0137	0.0565	0.0952
	<i>Fusión</i>	0.0360	0.0290	0.0255	0.0139	0.0552	0.0981
Intervalos	Intervalo temporal	0.0304	0.0251	0.0221	0.0114	0.0473	0.0824
	Distancia geográfica	0.0287	0.0238	0.0213	0.0107	0.0447	0.0795
Modelo	Sin Módulo TCN	0.0345	0.0265	0.0231	0.0132	0.0500	0.0867
	Sin Atributos	0.0335	0.0283	0.0249	0.0128	0.0536	0.0933
	Mejor modelo (MAX)	0.0340	0.0283	0.0251	0.0132	0.0542	0.0942
	Mejor modelo (AVG)	0.0386	0.0315	0.0275	0.0149	0.0597	0.1014

Yelp

Influencia de los atributos

Igual que en el caso anterior el intervalo temporal es superior al intervalo geográfico

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Comentarios	Utilidad	0.0328	0.0264	0.0231	0.0129	0.0508	0.0877
	Gracioso	0.0325	0.0264	0.0232	0.0125	0.0502	0.0875
	Interesante	0.0316	0.0269	0.0237	0.0121	0.0511	0.0890
	Fecha	0.0332	0.0271	0.0240	0.0128	0.0511	0.0911
	<i>Fusión</i>	0.0348	0.0295	0.0259	0.0132	0.0551	0.0961
Geografía	Dirección	0.0339	0.0276	0.0250	0.0134	0.0528	0.0938
	Ciudad	0.0343	0.0283	0.0249	0.0129	0.0535	0.0939
	Estado	0.0336	0.0270	0.0239	0.0129	0.0513	0.0903
	Código postal	0.0358	0.0294	0.0255	0.0137	0.0565	0.0952
	<i>Fusión</i>	0.0360	0.0290	0.0255	0.0139	0.0552	0.0981
Intervalos	Intervalo temporal	0.0304	0.0251	0.0221	0.0114	0.0473	0.0824
	Distancia geográfica	0.0287	0.0238	0.0213	0.0107	0.0447	0.0795
Modelo	Sin Módulo TCN	0.0345	0.0265	0.0231	0.0132	0.0500	0.0867
	Sin Atributos	0.0335	0.0283	0.0249	0.0128	0.0536	0.0933
	Mejor modelo (MAX)	0.0340	0.0283	0.0251	0.0132	0.0542	0.0942
	Mejor modelo (AVG)	0.0386	0.0315	0.0275	0.0149	0.0597	0.1014

Yelp

Influencia de los atributos

Modulo TCN es mejor

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Comentarios	Utilidad	0.0328	0.0264	0.0231	0.0129	0.0508	0.0877
	Gracioso	0.0325	0.0264	0.0232	0.0125	0.0502	0.0875
	Interesante	0.0316	0.0269	0.0237	0.0121	0.0511	0.0890
	Fecha	0.0332	0.0271	0.0240	0.0128	0.0511	0.0911
	<i>Fusión</i>	0.0348	0.0295	0.0259	0.0132	0.0551	0.0961
Geografía	Dirección	0.0339	0.0276	0.0250	0.0134	0.0528	0.0938
	Ciudad	0.0343	0.0283	0.0249	0.0129	0.0535	0.0939
	Estado	0.0336	0.0270	0.0239	0.0129	0.0513	0.0903
	Código postal	0.0358	0.0294	0.0255	0.0137	0.0565	0.0952
	<i>Fusión</i>	0.0360	0.0290	0.0255	0.0139	0.0552	0.0981
Intervalos	Intervalo temporal	0.0304	0.0251	0.0221	0.0114	0.0473	0.0824
	Distancia geográfica	0.0287	0.0238	0.0213	0.0107	0.0447	0.0795
Modelo	Sin Módulo TCN	0.0345	0.0265	0.0231	0.0132	0.0500	0.0867
	Sin Atributos	0.0335	0.0283	0.0249	0.0128	0.0536	0.0933
	Mejor modelo (MAX)	0.0340	0.0283	0.0251	0.0132	0.0542	0.0942
	Mejor modelo (AVG)	0.0386	0.0315	0.0275	0.0149	0.0597	0.1014

Yelp

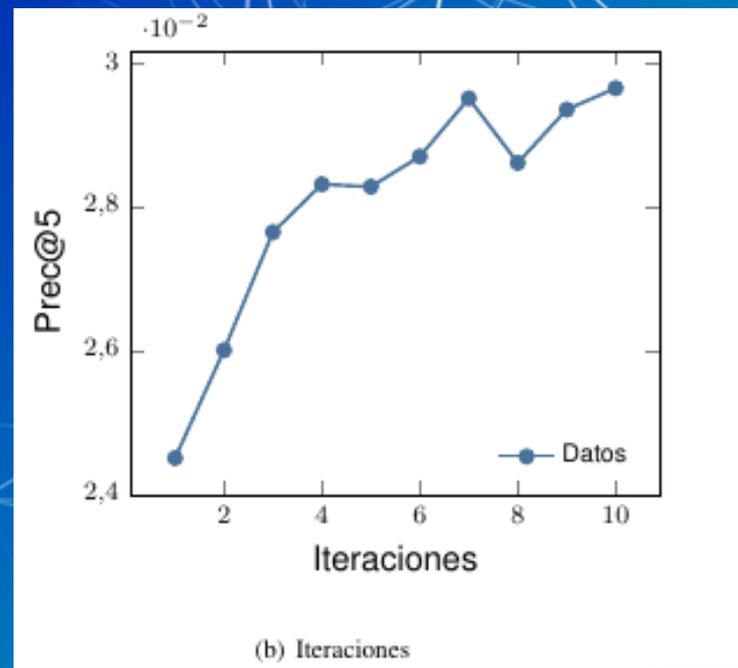
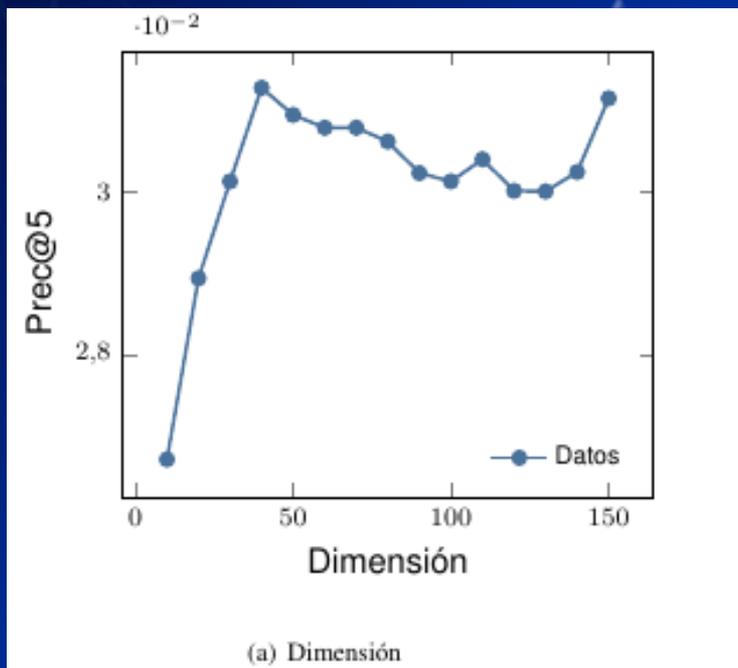
Influencia de los atributos

Conseguimos los mejores resultados con la operación de media

Tipo	Nombre	Prec@1	Prec@5	Prec@10	Recall@1	Recall@5	Recall@10
Comentarios	Utilidad	0.0328	0.0264	0.0231	0.0129	0.0508	0.0877
	Gracioso	0.0325	0.0264	0.0232	0.0125	0.0502	0.0875
	Interesante	0.0316	0.0269	0.0237	0.0121	0.0511	0.0890
	Fecha	0.0332	0.0271	0.0240	0.0128	0.0511	0.0911
	<i>Fusión</i>	0.0348	0.0295	0.0259	0.0132	0.0551	0.0961
Geografía	Dirección	0.0339	0.0276	0.0250	0.0134	0.0528	0.0938
	Ciudad	0.0343	0.0283	0.0249	0.0129	0.0535	0.0939
	Estado	0.0336	0.0270	0.0239	0.0129	0.0513	0.0903
	Código postal	0.0358	0.0294	0.0255	0.0137	0.0565	0.0952
	<i>Fusión</i>	0.0360	0.0290	0.0255	0.0139	0.0552	0.0981
Intervalos	Intervalo temporal	0.0304	0.0251	0.0221	0.0114	0.0473	0.0824
	Distancia geográfica	0.0287	0.0238	0.0213	0.0107	0.0447	0.0795
Modelo	Sin Módulo TCN	0.0345	0.0265	0.0231	0.0132	0.0500	0.0867
	Sin Atributos	0.0335	0.0283	0.0249	0.0128	0.0536	0.0933
	Mejor modelo (MAX)	0.0340	0.0283	0.0251	0.0132	0.0542	0.0942
	Mejor modelo (AVG)	0.0386	0.0315	0.0275	0.0149	0.0597	0.1014

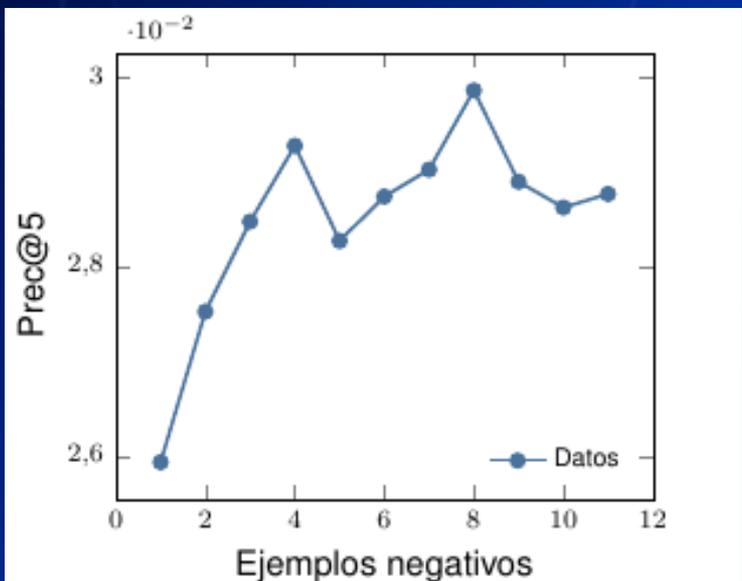
Yelp

Análisis de hiperparámetros

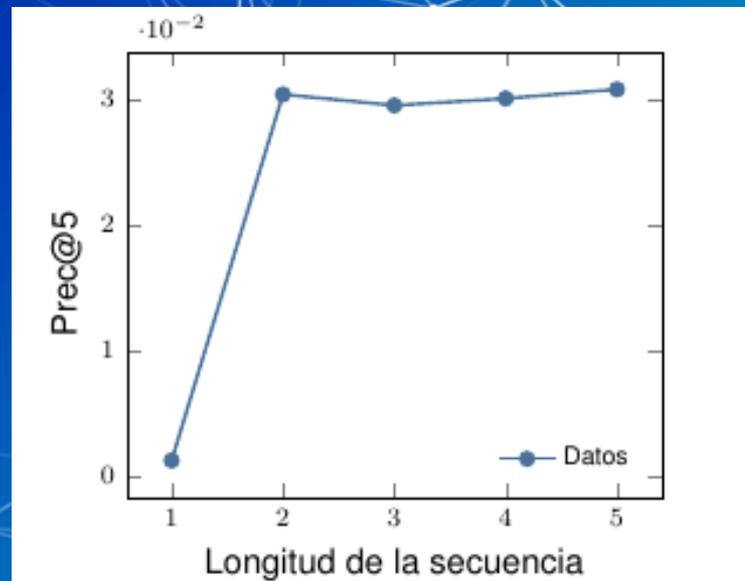


Yelp

Análisis de hiperparametros



(c) Ejemplos negativos



(d) Longitud

Índice

- Introducción
- Estado del arte
- Propuesta
- Diseño e implementación
- Experimentos y resultados
- **Conclusiones y trabajo futuro**

Conclusiones y trabajo futuro

- Las operaciones de máximo y media aplicadas a los vectores de embedding han dado buenos resultados teniendo un mayor rendimiento la operación de **media**.
- A medida que nuestro modelo tiene más atributos auxiliares mejor es el rendimiento que obtenemos.
- Existen atributos que aportan más valor que otros al modelo como los atributos geográficos o la fecha de cuando se realizó un viaje o un comentario.
- Los parámetros ideales de un modelo neuronal deben realizarse siempre que se pueda mediante barrido puesto que son vitales para obtener un buen rendimiento del modelo.
- Utilizar valores pre-entrenados para la inicialización de los embeddings parece favorecer siempre a los resultados.



Conclusiones y trabajo futuro

- Ampliar experimentos utilizando otros conjuntos de datos y métricas.
- Incorporar otros atributos como imágenes o el texto de las reviews en el modelo neuronal.
- El uso de redes neuronales recurrentes con mecanismos de atención.



**MUCHAS
GRACIAS!!**