Agenda

- anuncios varios
  - Tarea 3 entrega lunes 5 de Junio
- modelos de analitica (machine learning-ML) Supe
  - Redes Neuronales
    - Resumen Convolutional Neural Net
    - LSTM
      - Series de tiempo
      - Secuencias

Series de Tiempo

La predicción actual depende de los datos históricos :

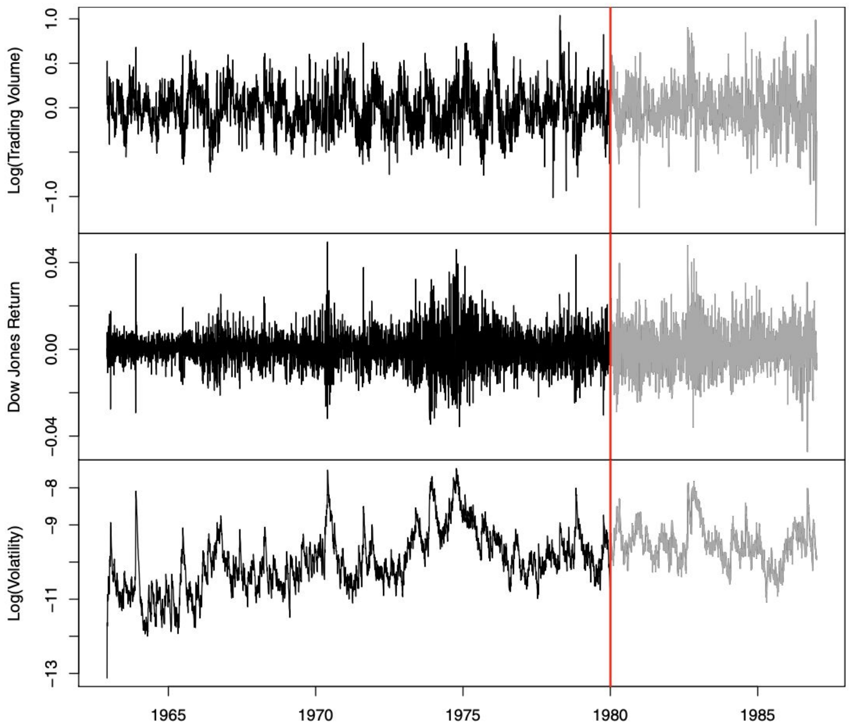


FIGURE 10.14. *Historical trading statistics from the New York Stock Exchange. Daily values of the normalized log trading volume, DJIA return, and log volatility are shown for a 24-year period from 1962–1986. We wish to predict trading volume on any day, given the history on all earlier days. To the left of the red bar (January 2, 1980) is training data, and to the right test data.*

$$X_1 = \begin{pmatrix} v_{t-L} \\ r_{t-L} \\ z_{t-L} \end{pmatrix}, \; X_2 = \begin{pmatrix} v_{t-L+1} \\ r_{t-L+1} \\ z_{t-L+1} \end{pmatrix}, \cdots$$

$$, X_L = \begin{pmatrix} v_{t-1} \\ r_{t-1} \\ z_{t-1} \end{pmatrix}, \quad \text{and } Y = v_t.$$

* Gareth James, An Introduction to Statistical Learning

Series de Tiempo

Redes que permiten aprender las relaciones en las secue

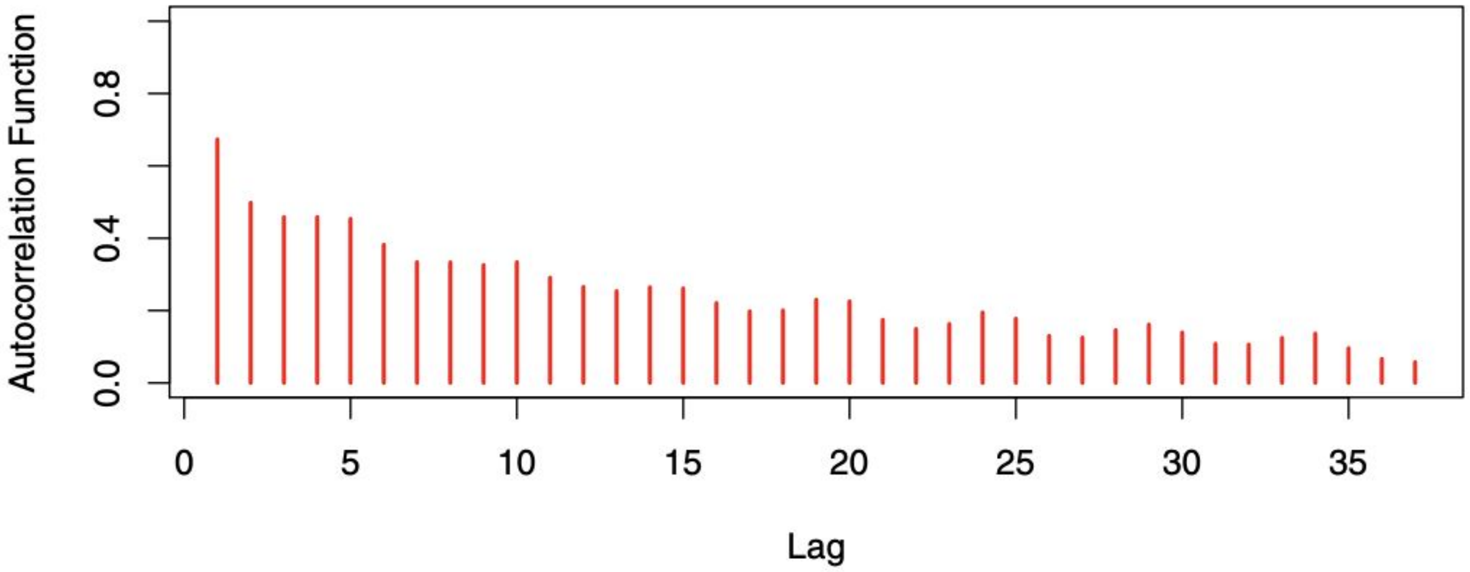


FIGURE 10.15. *The autocorrelation function for* `log_volume`. *We see that nearby values are fairly strongly correlated, with correlations above 0.2 as far as 20 days apart.*

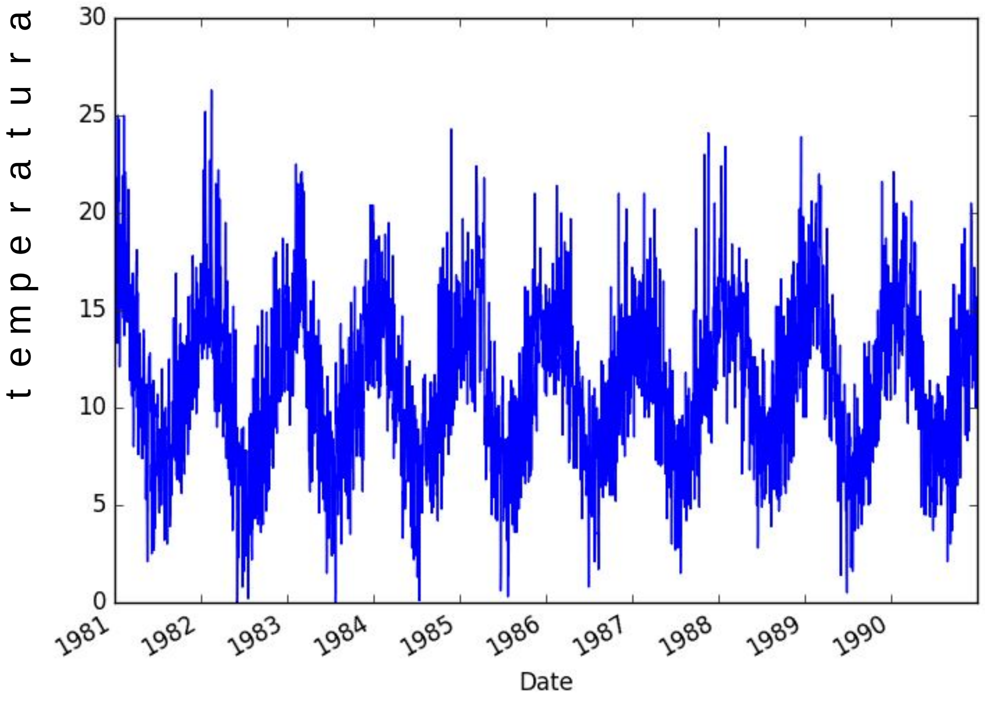
Es necesario defin

Window: numero de
Lag: Numero de obs
anteriores

* Gareth James, An Introduction to Statistical Learning

Series de Tiempo

Seasonality son los periodos de tiempo que tienen un pa

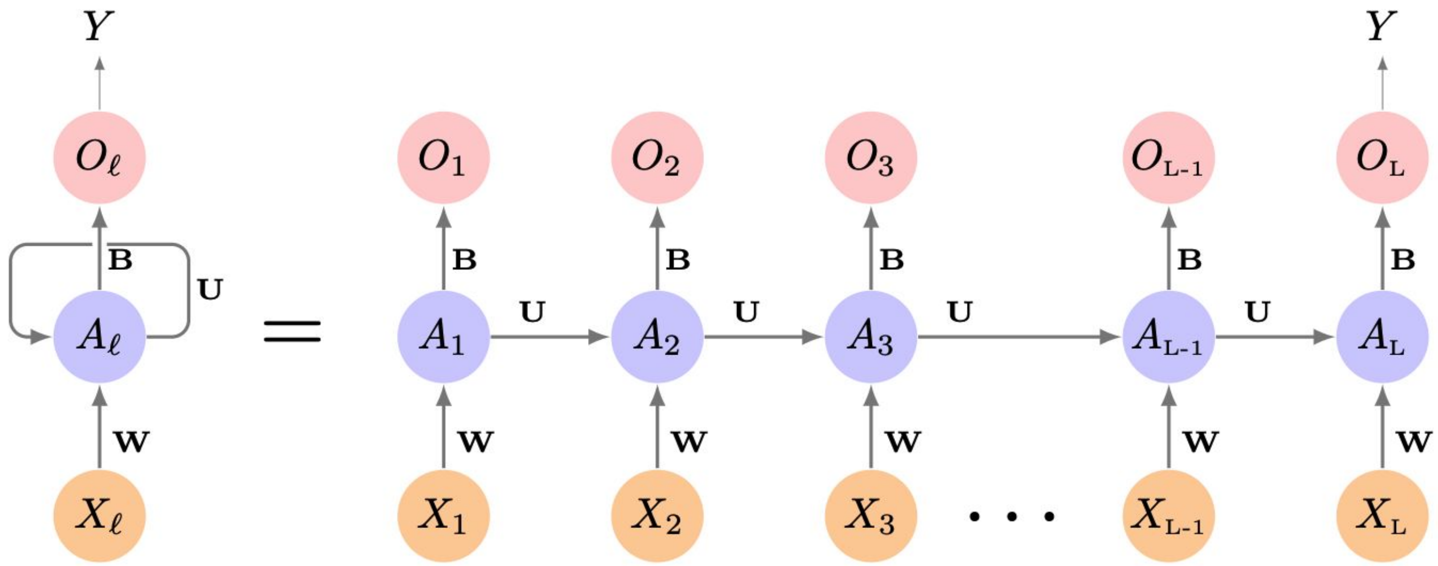


Periodos:
- diaria
- semanal
- mensual
- trimestral
- semestral
- anual

* Gareth James, An Introduction to Statistical Learning

Recurrent neural nets

Redes que permiten aprender las relaciones en las secue



ERROR

$$(Y - O_L)^2,$$

$$A_{\ell k} = g\left(w_{k0} + \sum_{j=1}^{p} w_{kj} X_{\ell j} + \sum_{s=1}^{K} u_{ks} A_{\ell-1,s}\right), \qquad O_\ell = \beta_0 + \sum_{k=1}^{K} \beta_k A_{\ell k}$$

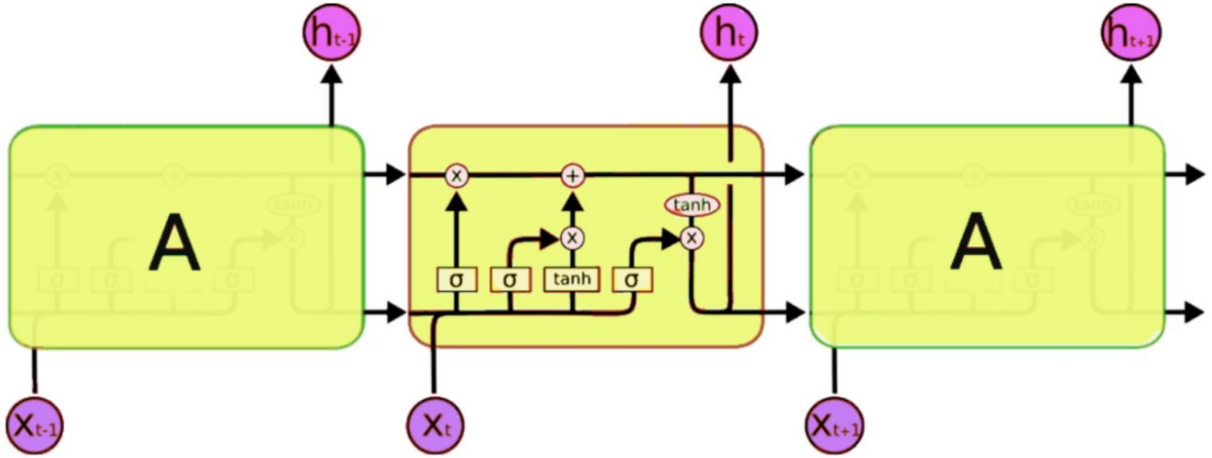
* Gareth James, An Introduction to Statistical Learning

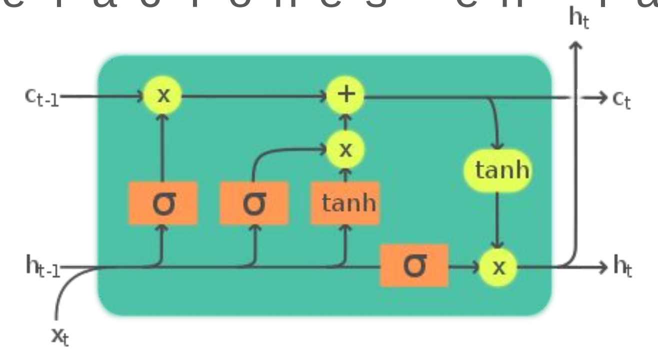
Agenda

- anuncios varios
    - Tarea 3 entrega lunes 5 de Junio
- modelos de analitica (machine learning-ML) Supe
    - Redes Neuronales
        - LSTM
            - Secuencias

Recurrent neural nets, LSTM

Redes que permiten aprender las relaciones en las secue





$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f)$$
$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i)$$
$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o)$$
$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$
$$h_t = o_t \odot \sigma_h(c_t)$$

http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf
Krizhevsky (2009) “Learning multiple layers of features from tiny images”, avai

* Gareth James, An Introduction to Statistical Learning

Recurrent neural nets, ejemplo

La secuencia de eventos es una variab

SISBEN Fecha de Registro

Nace

2011

Se muda a otro municipio

Se casa

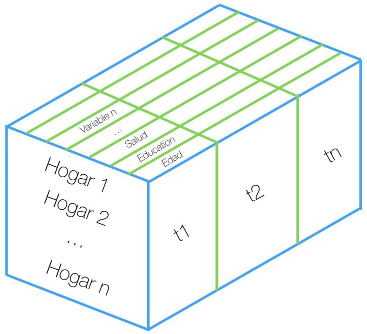
2016

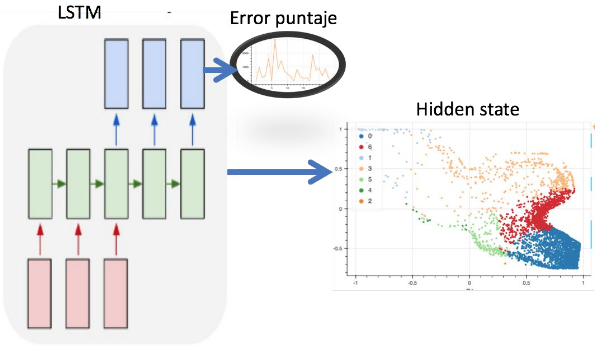
Puntaje SISBEN

Variables de Vivienda

LSTM, predecir anomalías

Salida, probabilidad de fraude





Ranking order

+ -

riesgo

t=11, variables sin procesar # 444.

El modelo usa dos capas con (100, 20) neuronas

http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf

Krizhevsky (2009) “Learning multiple layers of features from tiny images”, avai

* Gareth James, An Introduction to Statistical Learning

Recurrent neural nets, ejemplo

Ajuster en las series de tiempo

- Preparar datos
  - Stemming: Remover las terminaciones de las palabras para que sean
    oconjugación
    - Estudiant
    - Jugar
  - Lemmatization: Utilizar el origen de las palabra
    - was = be
    - were = be
    - are = be

http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf
Krizhevsky (2009) “Learning multiple layers of features from tiny images”, avai

* Gareth James, An Introduction to Statistical Learning

Recurrent neural nets, LSTM para

Redes que permiten aprender las relaciones en las secue

- one_hot: es un vector donde cada palabra es

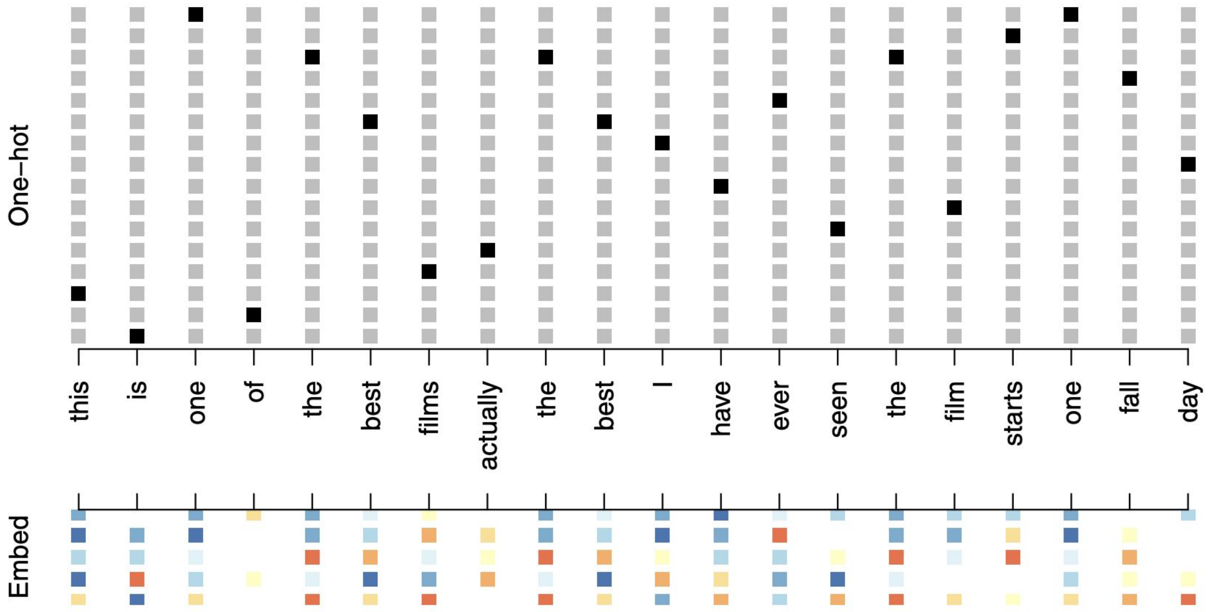


FIGURE 10.13. *Depiction of a sequence of 20 words representing a single document: one-hot encoded using a dictionary of 16 words (top panel) and embedded in an m-dimensional space with $m = 5$ (bottom panel).*

http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf

Krizhevsky (2009) "Learning multiple layers of features from tiny images", avai

* Gareth James, An Introduction to Statistical Learning

Recurrent neural nets, LSTM para

Ajuster en las series de tiempo

- Utilizar word2vec: Embedding para representar las supervisado.

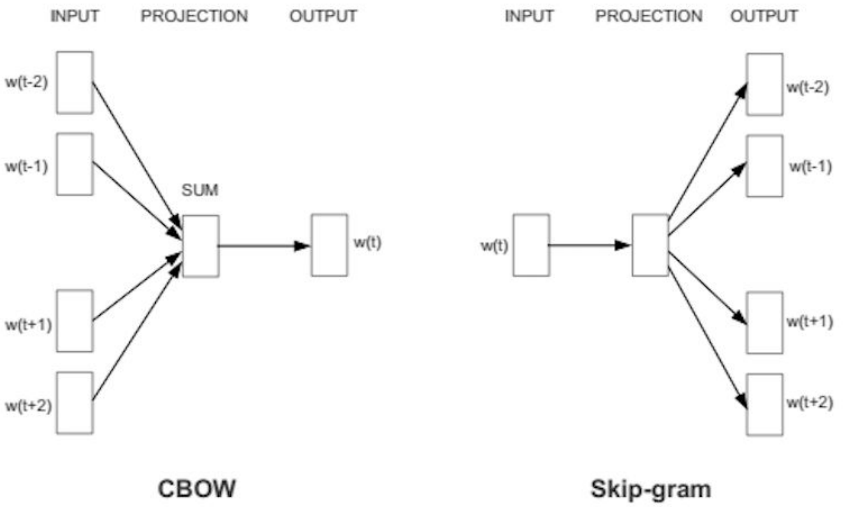
Instead of only looking two words before the target word, we can also look at two words after it.

Jay was hit by a ______ bus in...

by	a	red	bus	in

If we do this, the dataset we're virtually building and training the model against would look like this:

input 1	input 2	input 3	input 4	output
by	a	bus	in	red



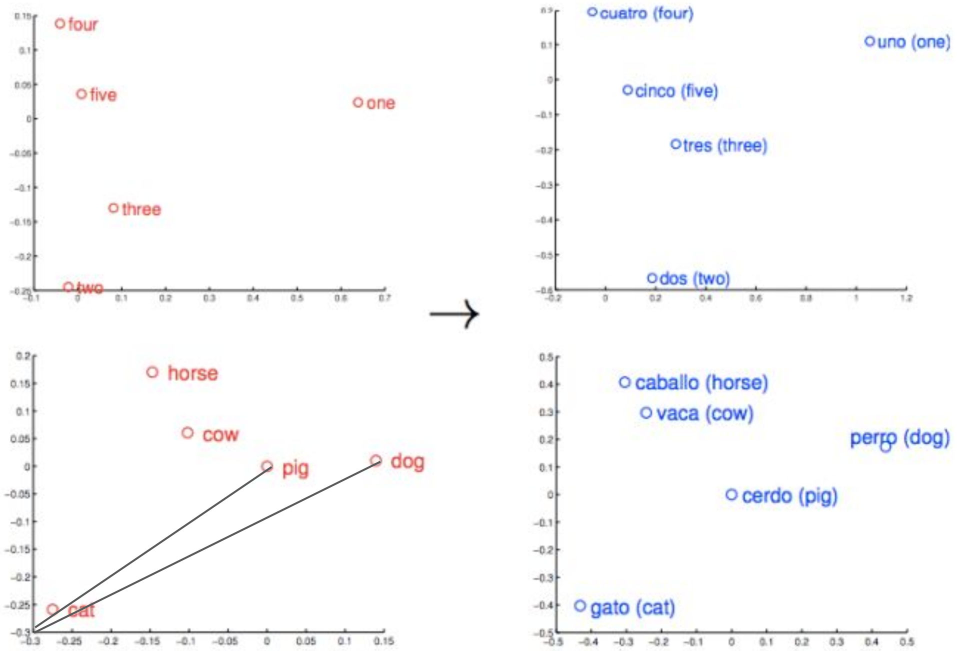
https://jalammar.github.io/illustrated-word2vec/
Word2vec paper https://arxiv.org/pdf/1301.3781.pdf

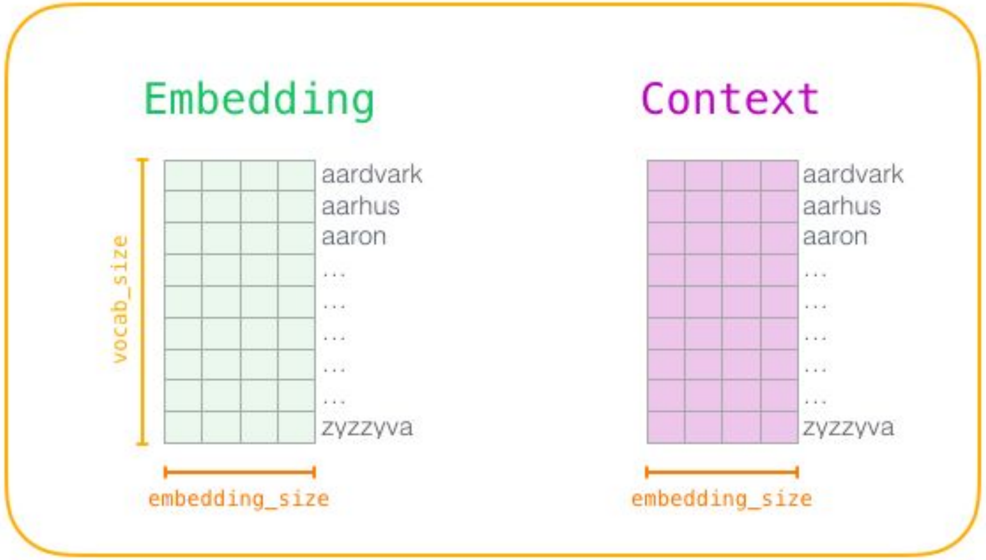
* Gareth James, An Introduction to Statistical Learning

Recurrent neural nets, LSTM para

Ajuster en las series de tiempo

- Utilizar word2vec: Embedding para representar las
  supervisado.





https://jalammar.github.io/illustrated-word2vec/
Wrod2vec paper https://arxiv.org/pdf/1301.3781.pdf

* Gareth James, An Introduction to Statistical Learning

Recurrent neural nets, ejemplo

Ajuster en las series de tiempo

- Definir window size: Tamaño de la secuencia
  Cuántas palabras se utilizaran para prede

http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf
Krizhevsky (2009) "Learning multiple layers of features from tiny images", avai

* Gareth James, An Introduction to Statistical Learning

Recurrent neural nets, ejemplo I

Ajuster en las series de tiempo

- Definir arquitectura: Cuantas capas y neuro

http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf
Krizhevsky (2009) “Learning multiple layers of features from tiny images”, avai
* Gareth James, An Introduction to Statistical Learning