

SPEECH TAGGING CON HMM's

Categorizar las partes de un discurso (verbos, adjetivos, etc.) es útil para modelos encargados de corregir la gramática o la ortografía.

Por simplicidad, utilizaremos 3 categorías gramaticales

- ↳ Noun: House, car, Mary, etc.
- ↳ Modal verb (Usado con otro verbo): can, would, should, etc.
- ↳ Verb: Run, see, jump, etc.

Data: Mary saw Jane <u>Jane</u> <u>saw</u> <u>Will</u> <u>N</u> <u>V</u> <u>N</u>	\Rightarrow	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th></th> <th style="background-color: #0070C0;"></th> <th style="background-color: #0070C0;"></th> </tr> <tr> <th style="background-color: #0070C0;"></th> <th style="background-color: #0070C0;">N</th> <th style="background-color: #0070C0;">V</th> </tr> </thead> <tbody> <tr> <td>Mary</td> <td>1</td> <td>0</td> </tr> <tr> <td>saw</td> <td>0</td> <td>2</td> </tr> <tr> <td>Jane</td> <td>1</td> <td>0</td> </tr> <tr> <td>Will</td> <td>1</td> <td>0</td> </tr> </tbody> </table>					N	V	Mary	1	0	saw	0	2	Jane	1	0	Will	1	0	\leftarrow Lookup Table
	N	V																			
Mary	1	0																			
saw	0	2																			
Jane	1	0																			
Will	1	0																			

Existen otros tipos de tablas como los bigrams donde se analiza la relación entre una palabra y su palabra vecina:

Ex.: "Mary will see Jane" \rightarrow

	N-M	M-V	V-N
Mary-will	1	0	0
will-see	0	1	0
see-Jane	0	0	1

Sin embargo, al presentarle a nuestro modelo un ejemplo que no se encuentre en la tabla, no tendría forma de añadir esta información extra en ella.

↳ Para solucionar este problema utilizamos los Hidden Markov Models:

Vamos a establecer 2 probabilidades:

↳ Transition prob.: Relación entre palabras.

↳ Emission prob.: Probabilidad de que una categoría grammatical corresponda a una determinada palabra.

Ex.:

Start tag						End tag
	N	N	M	V	N	

(S) Mary Jane can see Will (E)
 (S) Spot will see Mary (E) \Rightarrow
 (S) Will Jane spot Mary? (E)
 (S) Mary will pat Spot (E)

Emission probability

	N	M	V
Mary	4/9	0	0
Jane	2/9	0	0
Will	1/9	3/4	0
Spot	2/9	0	1/4
Can	0	1/4	0
See	0	0	1/2
Pat	0	0	1/4

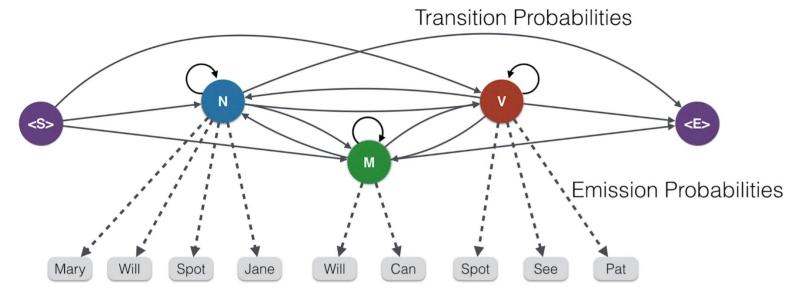
Dividimos cada valor entre la suma de los elementos de su columna

Transition probability

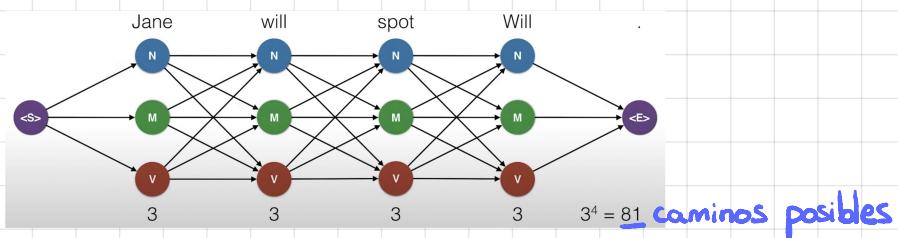
	N	M	V	(E)
(S)	3/4	1/4	0	0
N	1/4	1/3	1/9	7/9
M	1/4	0	3/4	0
V	1	0	0	0

Dividimos cada valor entre la suma de los elementos de su fila

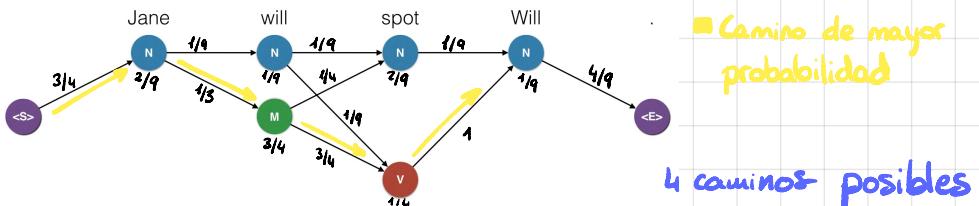
Finalmente, para formar el Hidden Markov Model combinaremos los gráficos resultantes de ambas probabilidades:



Si hacemos un diagrama de todas las posibilidades obtenemos un Trellis Diagram:



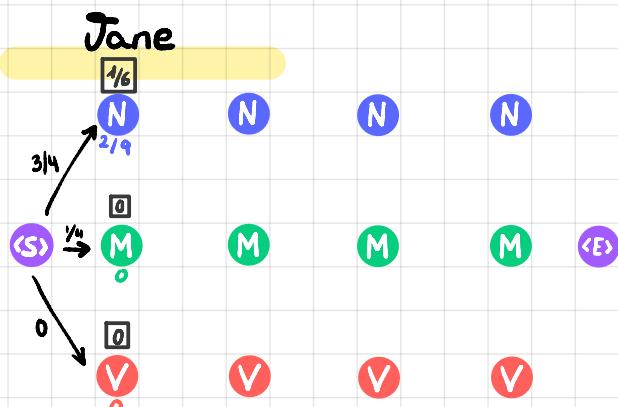
Y si borramos tanto las emission probabilities como las transition probabilities con valor 0 obtenemos el siguiente diagrama:



De esta forma (escogiendo el camino con mayor probabilidad) vemos como podemos etiquetar correctamente la oración "Jane will spot Will" como N-M-V-N.

Viterbi Algorithm:

Este algoritmo nos ayuda a encontrar el camino de mayor probabilidad.



	Emission		
	N	M	V
Hat	4/9	0	0
Jane	2/9	0	0
Will	1/9	3/4	0
Spot	2/9	0	1/4
Can	0	1/4	0
See	0	0	1/2
Pot	0	0	1/4

	Transition			
	N	M	V	<E>
<S>	3/4	1/4	0	0
N	1/4	1/3	1/9	4/9
M	1/4	0	3/4	0
V	1	0	0	0