

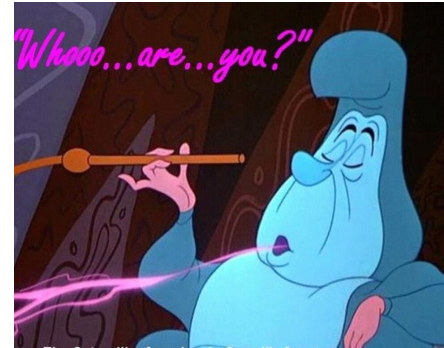


UNIVERSIDAD TECNOLÓGICA NACIONAL
FACULTAD REGIONAL CÓRDOBA

Diplomatura en Data Science Aplicada

Sobre mi...

- Me recibí de Ingeniero en Sistemas en la UTN FRC.
- Desde el 2011 hasta el 2018 fui investigador ad-honorem en el LIS.
- Me interesa mucho aprender cosas nuevas alrededor de data science.
- Apasionado python y buenas prácticas en el desarrollo de sw.
- Me gusta leer libros y practicar bonsai.



Sobre ustedes...



Objetivos del módulo

- Nociones básicas sobre NLP, estado del arte, y potenciales alcances.
- Análisis de corpus, comprender y entender el conjunto de datos.
- Técnicas básicas para tratar, transformar y generalizar textos.
- Conocimiento de distintos modelos, ventajas y potencial de cada uno.
- Aplicación práctica sobre un modelo de predicción.

Poniéndonos de acuerdo

- Este módulo vamos a trabajar con Python y Notebooks.
- Stack de datascience para Python.
- Trabajos prácticos grupales.
- No hay fechas duras para las entregas.



cartonero @flanchota · 18h
mi trabajo practico / el de mi amigo



888 5.904 41,9K



UNIVERSIDAD TECNOLÓGICA NACIONAL
FACULTAD REGIONAL CÓRDOBA

Introducción a PLN

Agenda

- História sobre NLP/PLN
- Estado del arte en NLP
- Motivaciones de la materia

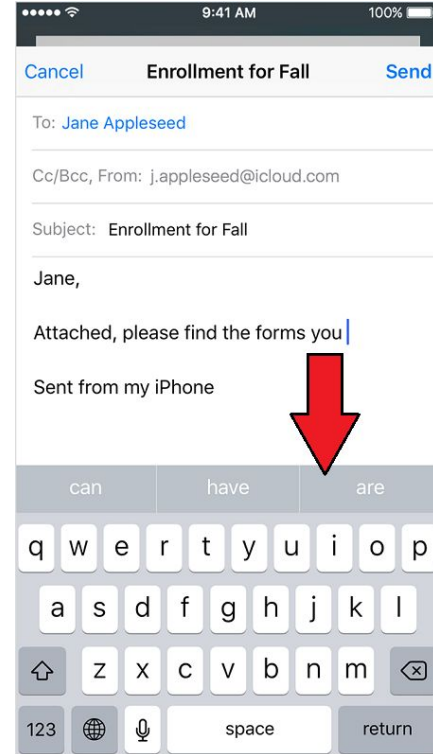
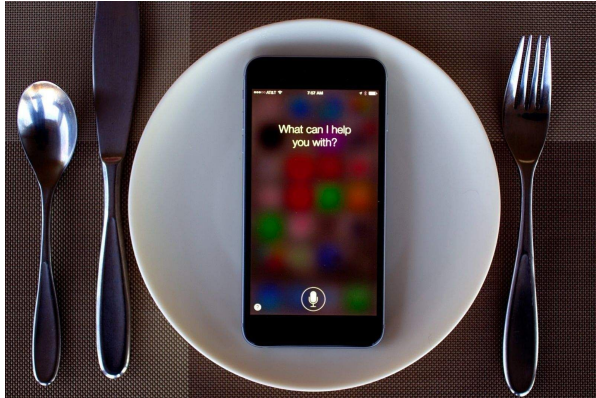


NLP - Respondiendo preguntas

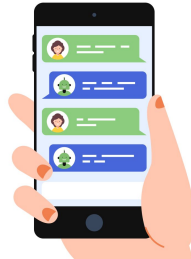


- El 16 de Febrero de 2011 Watson gana el concurso de preguntas y respuestas Jeopardy.
- Varios años antes, el 10 de Febrero de 1986 Deep blue gana un partido de ajedrez a Garry Kasparov

NLP: Traducciones automática



NLP: Chat bots e inteligencias artificiales



- GPT 2 -> 3 El algoritmo más “aterrador” de la historia.
- Más info del nuevo release [acá](#).

NLP: Chat bots e inteligencias artificiales



My experience
so far has been
fantastic!

POSITIVE



The product is
ok I guess

NEUTRAL



Your support team
is useless

NEGATIVE

Entonces... ¿Por qué no conquistamos el mundo?

- Detección de spam
- Part-of-speech tagging y Named entity recognition
- Traducciones a decenas de idiomas
- Sentiment analysis



- La ambigüedad propia en el lenguaje.
- Evolución en el lenguaje, retweet, friendzone, broman, broany.
- Modismos propios de cada idioma ([video](#)).
- Idioma no estándar, (ve vo ve) en redes sociales.



Consejos...

Sin importar si es este módulo o cual:

- Busca leer papes o investigaciones relacionado al campo de interés. Este lugar es un buen punto de partida: <https://arxiv.org/>
- Aprender algunas palabras técnicas relacionado al tópico (tech lingo).
- Participar de competencias en [Kaggle](#) y aprender de la ***kaggle experience***.
- Realiza pequeñas pruebas en tu tiempo libre. Si no tenés datos hay muchas fuentes con corpus en el caso de NLP o Kaggle tiene mucho de todo.

