

City University of New York (CUNY)

CUNY Academic Works

Open Educational Resources

City College of New York

2018

Lecture: Intro to Data Science - ML 3 - Week Eleven

Grant Long
CUNY City College

NYC Tech-in-Residence Corps

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/cc_oers/187

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).
Contact: AcademicWorks@cuny.edu

City College, Fall 2018

Intro to Data Science

Week 11: NLP, Text as Data, and Bayes Rule

November 19, 2018

Today's Agenda

1. What is NLP?
2. Bag of Words
Strategies
3. Bayesian Analysis
4. Tools and Topics in
NLP

A reminder:

Project updates due next Monday, November 26!

What is NLP?

Natural language processing (NLP) is a field of computer science, artificial intelligence, and computational linguistics concerned with the interactions between computers and human (natural) languages.

1. Identify the structure and meaning of words, sentences, texts and conversations.
2. Deep understanding of broad language.

Source.

NLP Applications: Quora Questions

Quora Home Answer Spaces Notifications Search Quora Add Question or Link

Neurolinguistic Programming Word Definitions, Terminology, and Jargon


What is NLP?

Answer Follow · 15 Request 1+ [Social Share Icons]

Ad by SPYSCAPE

The one thing you should do when in New York City.
New York's spy museum is open. Go on missions to test your hacking, deception, and espionage skills.
Learn more at spyscape.com

19 Answers

 **Phil Callaghan, Author of 'Hypnotic Conversations'**
Answered Mar 7, 2017 · Upvoted by Michael DeBusk, Society of NLP-trained since 1993

Given the category for this question is Neuro-Linguistic Programming, I'm going to assume that they're asking for some insight into that and not Natural Language Processing.

Neuro-Linguistic Programming is a way of looking at the world (*an attitude and a methodology*, to quote the originators' description), based on the work of Richard Bandler and John Grinder.

Drawing upon distinctions from Behavioural Psychology and Transformational Grammar, it aims to provide insight into our thinking processes.

There's more on Quora...

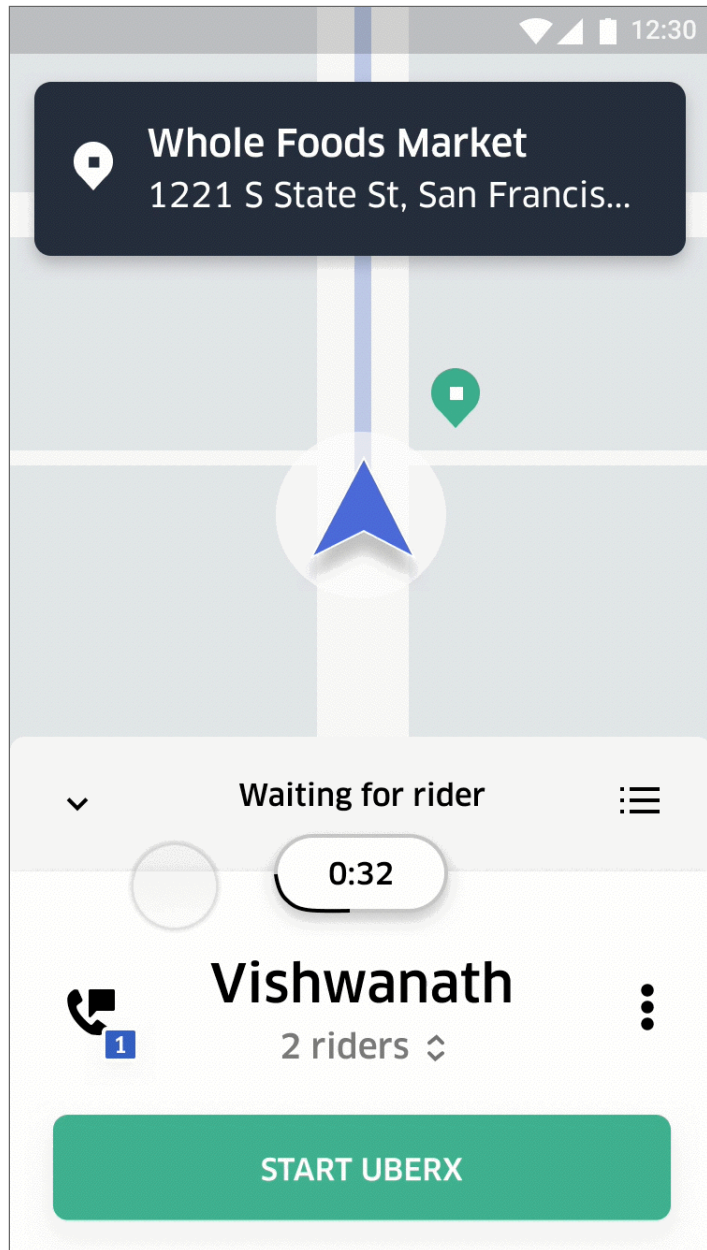
Pick new people and topics to follow and see the best answers on Quora.

[Update Your Interests](#)

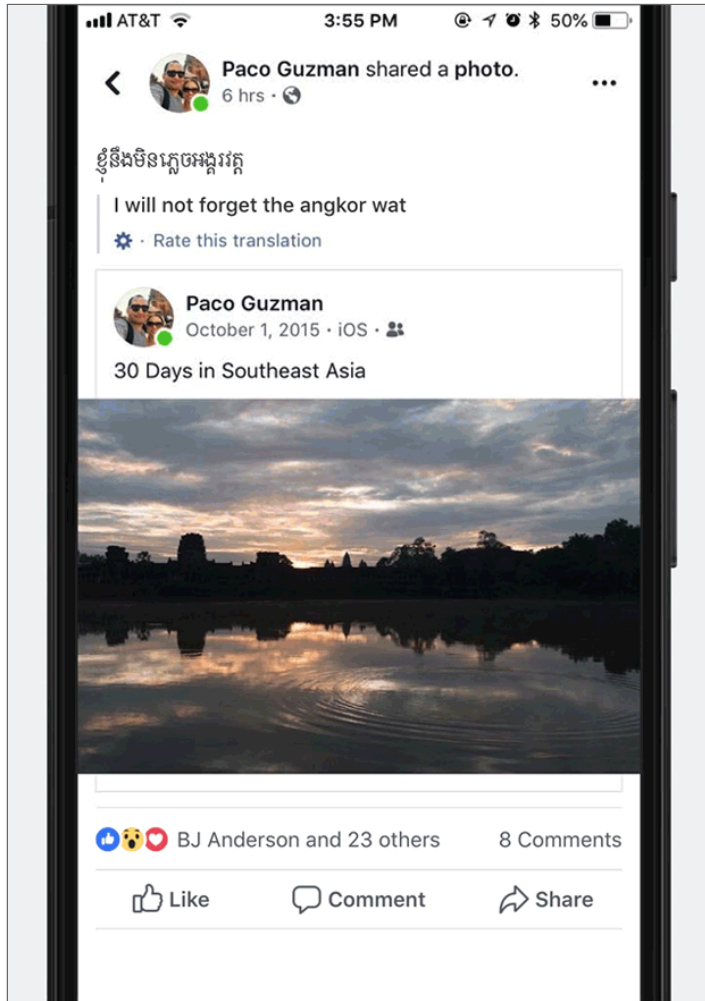
Related Questions

- Does neuro-linguistic programming really help, or does it have a placebo effect? If it does help, where can I get books, documents, or any kin...
- Does NLP change one's life?
- What is NLP training?
- What are the main differences between NLP for Chinese vs NLP for English?
- What is the best way to learn NLP?
- Can being aware of NLP indicators short-circuit NLP's accuracy and reliability?
- Is NLP useless?
- Is NLP manipulative?
- Who are the best certified trainers of NLP in India?
- Does Richard Bandler's NLP work?

NLP Applications: Uber One Click Chat



NLP Applications: Facebook Translation



Who's Excited?

More workmanlike than magical, "Fantastic Beasts: The Crimes of Grindelwald" nevertheless feels like an upgrade from its predecessor, one that adds star power, introduces key characters and lays the foundation for a genuine "Wizarding World" franchise. To call J.K. Rowling's mythology-heavy plot dense would be an understatement, but the film has enough epic heft to feel like a genuine blockbuster.

The labored second chapter in J.K. Rowling's Harry Potter spinoff series is as cumbersome as its title. "Fantastic Beasts: The Crimes of Grindelwald" is a gangly, overly complicated snooze, a rudderless, magic-free visit into Rowling's world of wizards and wizarding. Even the beasts aren't all that fantastic, and the visual effects aren't either.

More workmanlike than magical, "Fantastic Beasts: The Crimes of Grindelwald" nevertheless feels like an upgrade from its predecessor, one that adds star power, introduces key characters and lays the foundation for a genuine "Wizarding World" franchise. To call J.K. Rowling's mythology-heavy plot dense would be an understatement, but the film has enough epic heft to feel like a genuine blockbuster.

The labored second chapter in J.K. Rowling's Harry Potter spinoff series is as cumbersome as its title. "Fantastic Beasts: The Crimes of Grindelwald" is a gangly, overly complicated snooze, a rudderless, magic-free visit into Rowling's world of wizards and wizarding. Even the beasts aren't all that fantastic, and the visual effects aren't either.

How do we turn this into a model?

Bag of Words Strategies

Map words

Even the beasts aren't all that fantastic
5 1 3 9 11 2 10

to feature vectors.

| | | | | | | | | | | | | |
|-----------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|-----------|-----------|
| Vector | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

Bag of Words Strategies

Advantages:

- Easy to calculate
- Vectors are interpretable

Disadvantages:

- Word order often contains important context!
- Related: Interactions are hard (but not impossible) to model

Data Science Models

$$f(\text{features}) = \text{target}$$

~~Bey's~~ Bayesian Analysis

Bayes's Rules

The diagram shows the Bayes' Rule formula $P(c | x) = \frac{P(x | c)P(c)}{P(x)}$ with four labels and arrows: 'Likelihood' points to $P(x | c)$, 'Class Prior Probability' points to $P(c)$, 'Posterior Probability' points to $P(c | x)$, and 'Predictor Prior Probability' points to $P(x)$.

$$P(c | x) = \frac{P(x | c)P(c)}{P(x)}$$

Posterior Probability

Predictor Prior Probability

$$P(c | \mathbf{X}) = P(x_1 | c) \times P(x_2 | c) \times \cdots \times P(x_n | c) \times P(c)$$

Know this and how to apply it before any interview. It can get tricky.

Bayes's Rule and Text

$$P(\text{positive}|\text{words}) = \frac{P(\text{words}|\text{positive})P(\text{positive})}{P(\text{words})}$$

From Bayes's Rule to Naive Bayes

How to calculate?

$$P(\text{words}) = P(\text{"Even"} \cap \text{"the"} \cap \text{"beasts"})$$

Make some assumptions.

$$P(\text{words}) = P(\text{"Even"}) \times P(\text{"the"}) \times P(\text{"beasts"})$$

What Makes Naive Bayes *Naive*?

The conditional independence assumption: For conditional independence to hold true, we're assuming no words are more likely to appear with each other than any others.

Examples: "hot" and "dog", "ball" and "game", harry" and "potter", "computer" and "science"

Tools and Topics in NLP

Other Applications of Text Analysis:

- Part of Speech Tagging
- Named Entity Recognition
- Dependency Parsing
- Topic Modeling

For more, check out the Stanford **parser** and **named entity recognizer** and this **interactive topic modeling explorer**.

Helpful Open Source Tools

- **Sci-kit Learn**: Has helpful tools for featurization and basic text analysis.
- **NLTK**: Python package offering a wide array of functionality, including sentiment analysis, part of speech tagging, and more.
- **Gensim**: Python package offering a specialized tools in topic modeling and a few other domains.
- **Stanford Core NLP**: Java library offering a wide array of functionality *and state of the art* performance. (Easy interface with python available through NLTK).

Let's work some magic.

Assignment 7: Due Monday, December 3 by 6:30pm

DataCamp's Deep Learning in Python

- The course should appear collectively as assignment within your existing DataCamp account.
- Each section will appear separately and will be worth one point toward the total grade for the homework, plus an additional point for overall effort.
- Course claims to take 4 hours - as always, use your time wisely.