Natural Language Processing Introduction
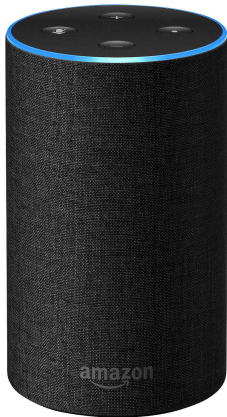
Natural Language Processing: Jordan Boyd-Graber

University of Maryland

00A

Roadmap

By the end of this class you should . . .

- Be able to give examples of where NLP is used
- Understand the workings of the course
- Know me and each other a little better

THINK

WATSON

$0	$3,600	-$200
KEN	WATSON	BRAD

Robert De Niro		84%
Chazz Palminteri		20%
Joe Pesci		14%

vary → **very**

The word **vary** doesn't s...
Consider replacing it

Google Übersetzer

Sofortübersetzung deaktivieren

| Englisch | Französisch | Deutsch | Sprache erkennen | ▾ | | Deutsch | Englisch | Chinesisch (vereinfacht) | ▾ | Übersetzen |

Da liegt der Hase im Pfeffer   ×   胡椒里有兔子

Computational Linguistics Applications

Machine Learning is Doing Great!

- Can drive a million miles without an accident
- Can beat any living chess player

Machine Learning is Doing Great!

- Can drive a million miles without an accident
- Can beat any living chess player
- Automated call center vs. five-year old?

Machine Learning is Doing Great!

- Can drive a million miles without an accident
- Can beat any living chess player
- Automated call center vs. five-year old?
- We'll learn why we're so far away

What's Natural Language Processing

- Computational approaches to understand, generate, and process natural language
- Cross-discipline
    - Computer science: implement algorithms
    - Linguistics: develop theory / data
    - Statistics: learn patterns from data
    - Experts in specific languages: get a computer to handle a new language
    - Psychologists: how does our brain process language
    - Sociologists: how do social constraints change how we process language

Our focus: question answering

We'll build a state-of-the-art question answering system

- Find related articles (information retrieval)
- Classifiers to identify relevant properties (machine learning)
- Read related source material (natural language understanding)
- Decide when to trust answers (sequence learning / calibration)

Our focus: question answering

We'll build a state-of-the-art question answering system

- Find related articles (information retrieval)
- Classifiers to identify relevant properties (machine learning)
- Read related source material (natural language understanding)
- Decide when to trust answers (sequence learning / calibration)
- Live competition with human players

What you need for this course

- Interactive classroom: come ready with questions and participate in exercises
- Helps to have a laptop to bring to class
- Math background
  - Will ask you to manipulate equations
  - Will expect you to be able to do basic derivatives
  - Work with functions like exponentiation and logs
  - Probability: review online (hugely important)
- Computer / programming skills
  - You will need to write python programs
  - You will need to interact with a Unix command line
  - You will need to interact with data files

Course reading
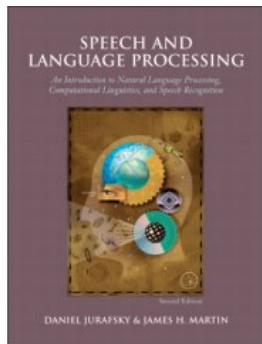
- We will provide reading assignments, mostly from the book. (Read them **before** associated class.)
- The reading will cover more than we cover in class.
- Free online

Before class . . .

- Install Python / nltk on your computer
- Come to class ready to work through some Python examples
- Also look at homework and see if you have any questions