



אוניברסיטת בר-אילן
הפקולטה למשפטים



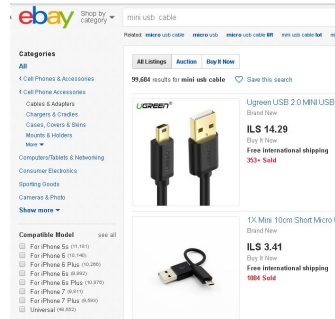
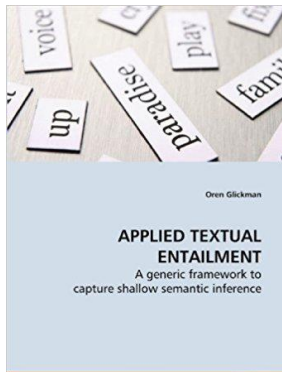
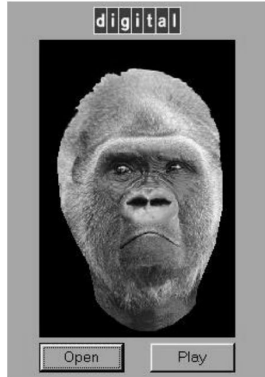
Data Science

for Legislation Studies

Dr. Oren Glickman

The Data Science Institute at Bar-Ilan University

About Myself



The Data Science Institute

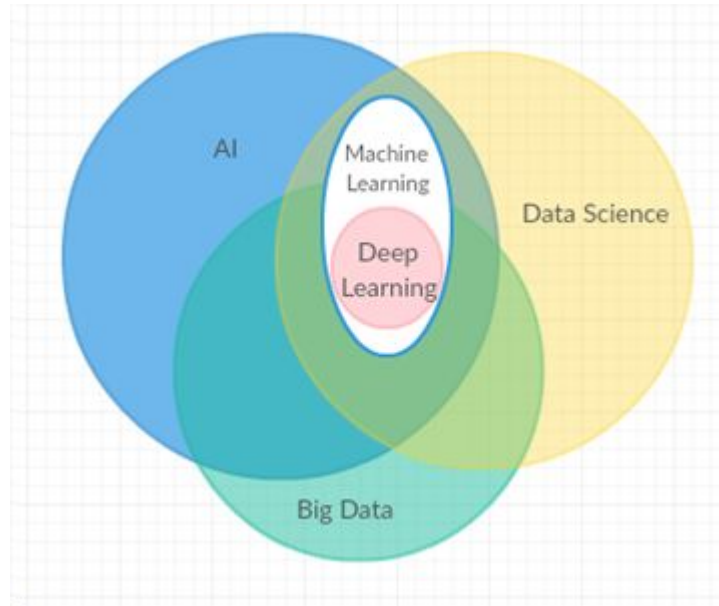
The DSI Mission:

Boost Core Data Science and Data Driven Research
across Bar-Ilan University



Terminology

Data Science vs. AI, Big Data, Machine Learning, Deep Learning



From Descriptive to Prescriptive



Data Sources

- Published laws and regulations
- Draft bills
- Transcripts of parliamentary debates and committees sessions
- Reports (state audit, committee reports, ...)
- ...

Methods

Applicable DS Methods for Legislation Studies:

- NLP/Text Analysis
- Networks/Graphs
- Predictive Analytics

Less relevant:

- Image Processing, Speech Processing, Time Series, Anomaly detection, ...

Research Examples

- Prediction of bill survival
- Automatic identification of draft bill changes
- Evaluate influence of interest groups
 - e.g. [Quantifying the influence of the tobacco industry](#) (Costa et al., 2014)
- Network Analysis
 - Cosponsorship networks
- Many more

NLP/Text Methods

Problems:

- Unstructured
- Infeasible to handle large volumes of textual data manually
- AI:
 - Language Ambiguity
 - Requires Natural Language Understanding
 - World and Domain Knowledge
 - Common Sense

Goal:

Effectively Turn text data into **high quality information**

NLP Methods

- Topic modeling / document clustering
- Text categorization
- Named entity extraction
- Relationship extraction
- Sentiment and opinion analysis
- Word Embeddings
- ...

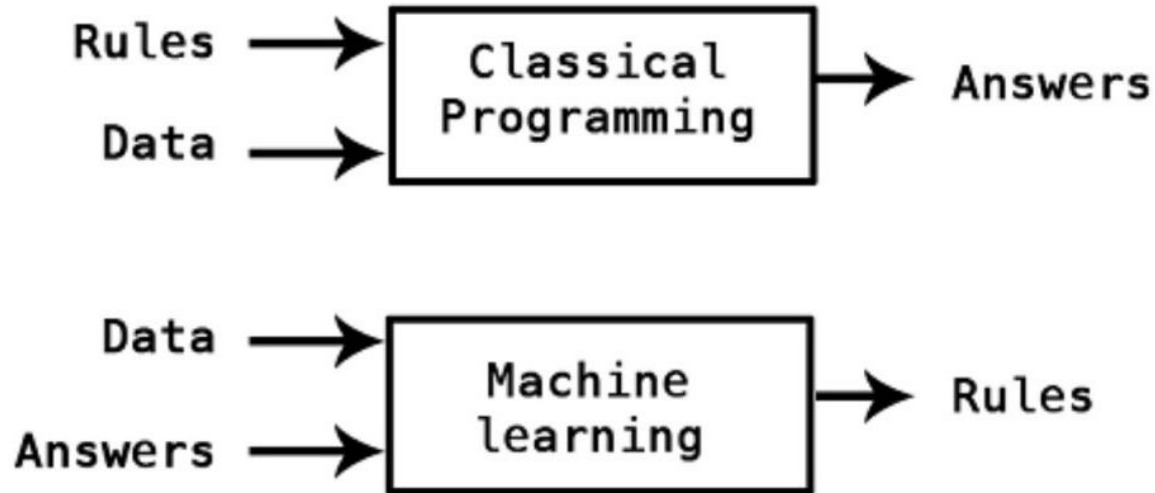
NLP Example: Extraction of Court Citations

- MS CS Project (Prof. Ido Dagan and Prof. Oren Perez)
- Automatically identify and extract academic citations in supreme court verdicts
- Applied various NLP algorithms using annotated dataset as supervised examples for learning
- **Demo:**
 - <http://35.225.98.72/link>
 - Example URL -
`https://supremedecisions.court.gov.il/Home/Download?path=HebrewVerdicts%5C15%5C480%5C012%5Ct13&fileName=15012480.T13&type=2`

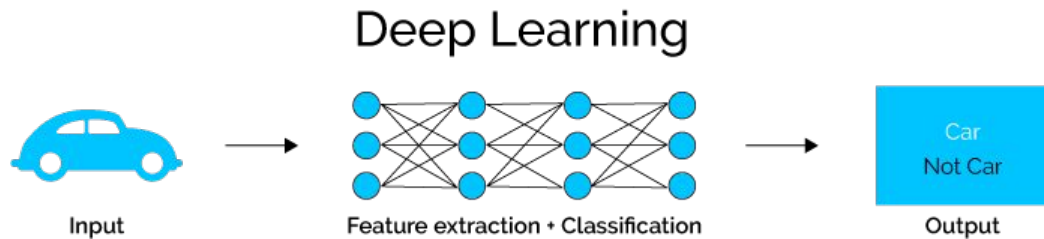
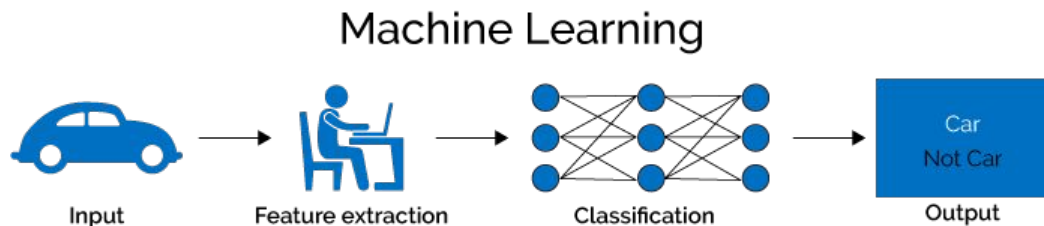
NLP Example:

- Explore Legislature - Supreme Court dialogue (I. Bar-Siman-Tov)
- In committees' transcripts, we are interested in constitutional discussions, references to past court decisions or to possible future rulings
- Can be manually annotated - but it's labor intensive, the texts of interest are rare
- Given some manual input, a computer algorithm can be trained to
 - Fully annotate more texts, but with limited accuracy (so far)
 - Find discussions passages where most of the texts of interest occur, saving most of the labor

Predictive Modeling in a nutshell



Deep Learning in a nutshell



Predictive Analytics Example

<http://aclweb.org/anthology/N12-1097>

Textual Predictors of Bill Survival in Congressional Committees

Combines non-textual fields such as committee, sponsor with the bills' content to predict its survival

Predictive Analytics Example

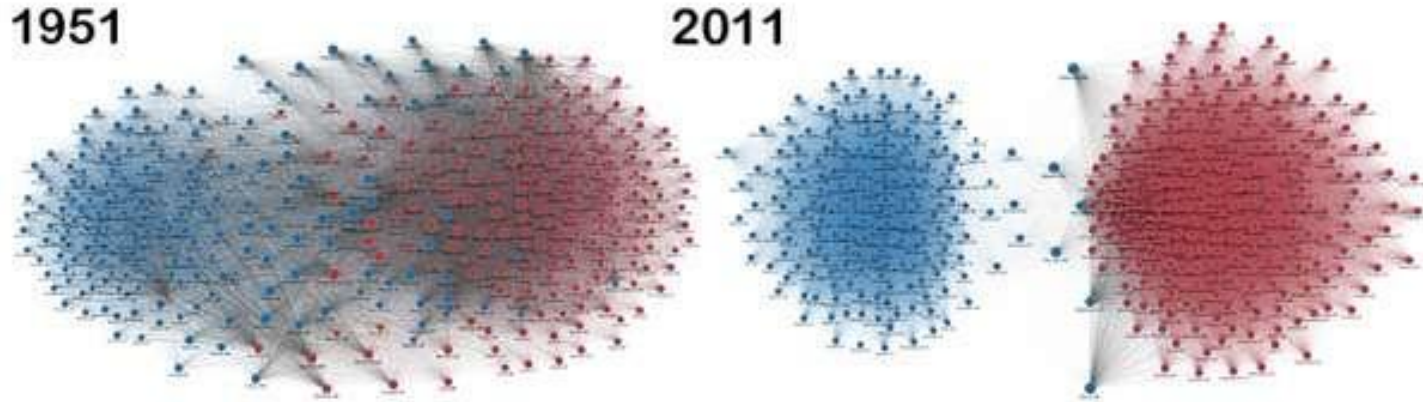
- Prediction of factory environmental pollution incidents
- Data from Ministry of Environmental Protection reports
- Goal:
 - Efficient use of limited enforcement and supervision resources
 - Avoid human biases



Network Analysis

- A network (or a graph) consists of nodes and edges connecting them
 - Directed, weighted, bipartite....
- General questions about networks/graphs:
 - Which nodes are most central to the network?
 - What are the Network properties of the graph - density, out vs. in degrees, connectivity?
 - How can the network be divided to sub-networks?
- In the study of legislation:
 - Nodes can represent legislators, parties, laws, institutions...
 - Edges can represent voting agreement, co-sponsorship of laws, similarities between texts...

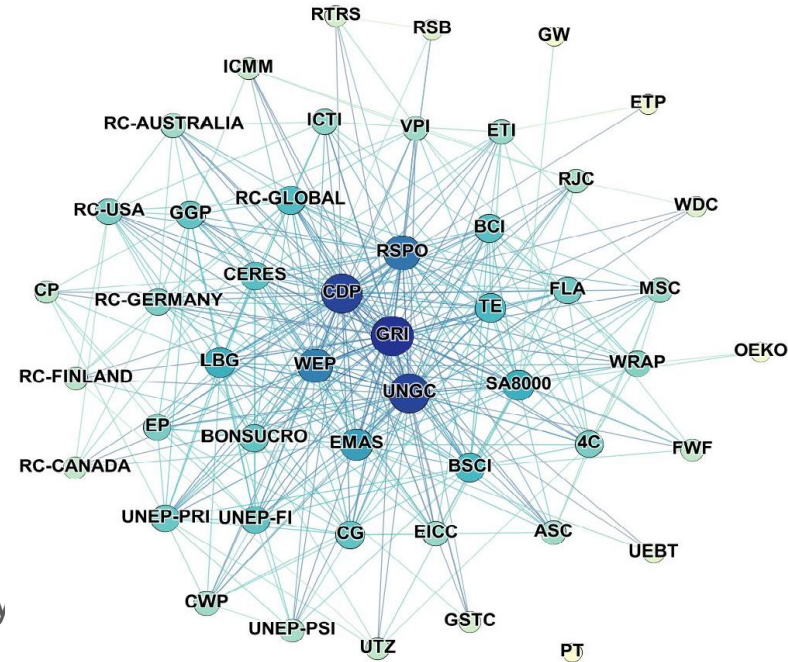
Network Analysis



Taken from: Andris et al., The Rise of Partisanship and Super-Cooperators in the U.S. House of Representatives, 2015

Example: Network Analysis of the Multinational Regulation

- 61 transnational CSR schemes and 31,987 firms
- Edge between 2 CSRs if share a common firm
- *Perez, Oren and Cohen, Reuven and Schreiber, Nir, **Governance through Global Networks and Corporate Signaling** (October 13, 2018). *Regulation and Governance* (2019). <https://ssrn.com/abstract=3265793>*
- Show that Membership in multiple schemes and the firms' position in the CSR-schemes network constitute credible predictors of their sustainability performance



Thank You