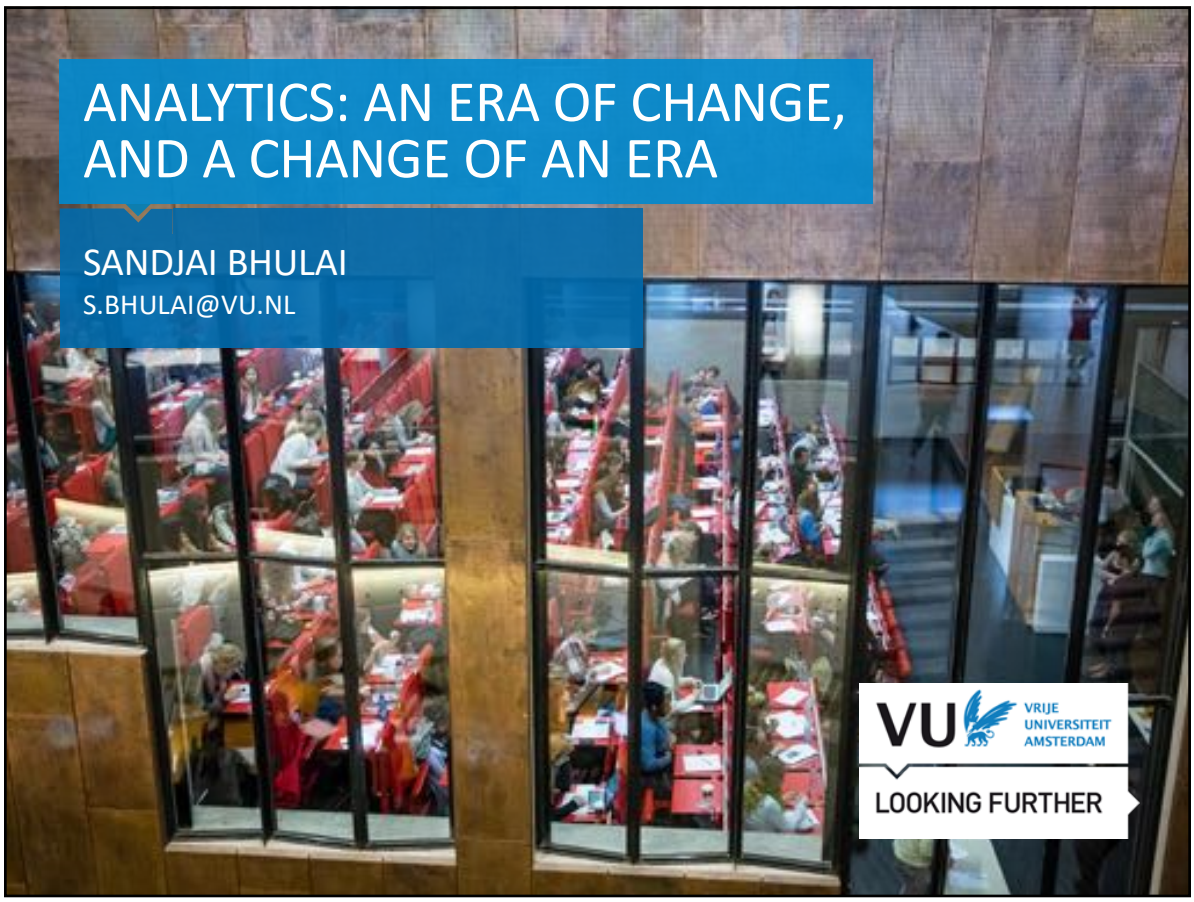
ANALYTICS: AN ERA OF CHANGE, AND A CHANGE OF AN ERA

SANDJAI BHULAI
S.BHULAI@VU.NL



1

INTRODUCTION

Prof. dr. S. (Sandjai) Bhulai
Full Professor of Business Analytics
Faculty of Science
Vrije Universiteit Amsterdam
Department: Mathematics
s.bhulai@vu.nl

Business Analytics: Optimization of Business Processes

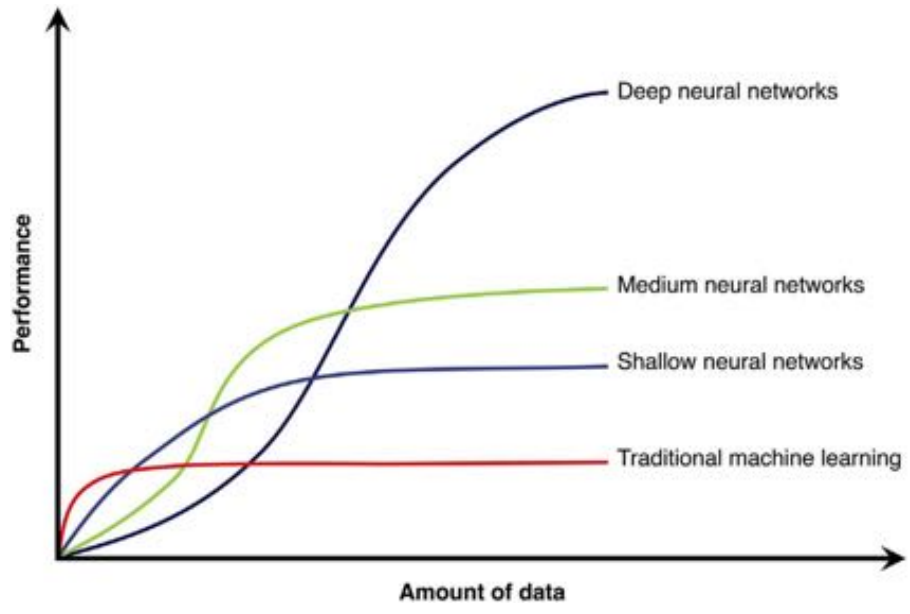
Research / Education:
Decision making under uncertainty
Control of complex high-dimensional systems
Stochastic optimization
Data science / machine learning

VU

2

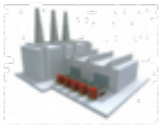
VOLUME OF DATA

VOLUME OF DATA



VARIETY OF DATA

- Structured data
- Unstructured data

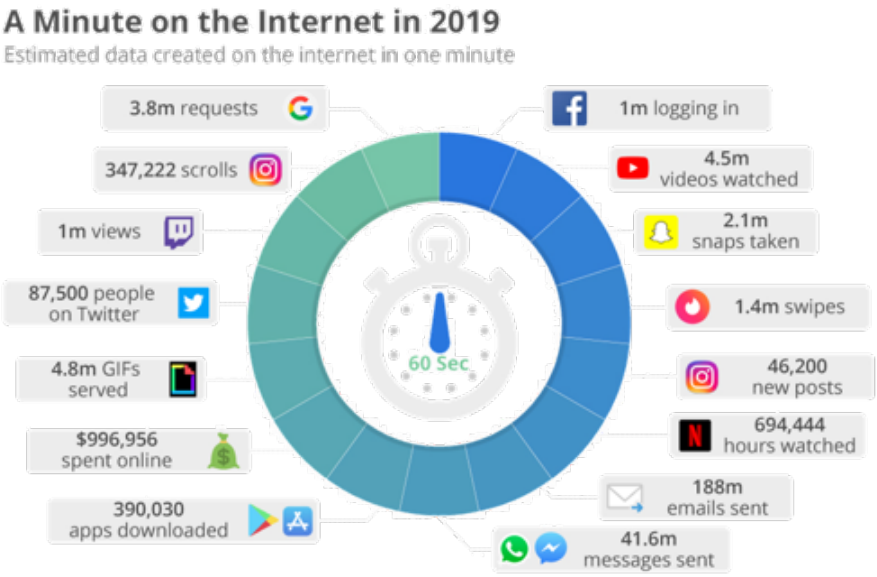


VU

VARIETY OF DATA

VU

VELOCITY OF DATA

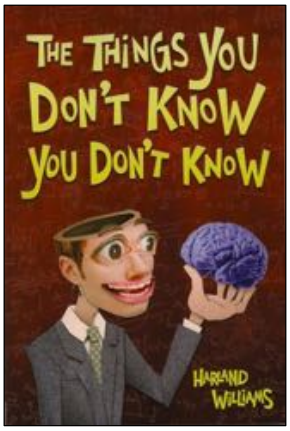


VELOCITY OF DATA

BUSINESS ANALYTICS

Big data is more than simply a matter of size: it is also about creating insights!

Discovering what you don't know you don't know……



BUSINESS ANALYTICS

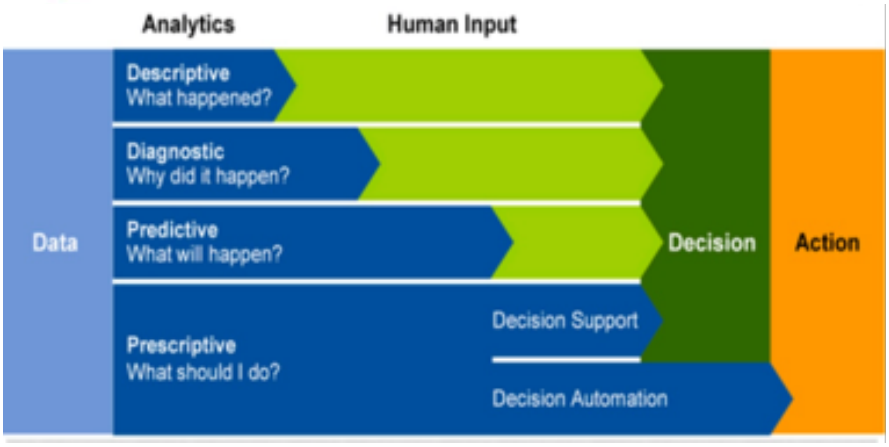
Big data is more than simply a matter of size: it is also about creating insights!

Business Analytics

- Scientific process of transforming data into insight for making better decisions
- Used for data-driven decision making, which is often seen as more objective than other alternatives for decision making

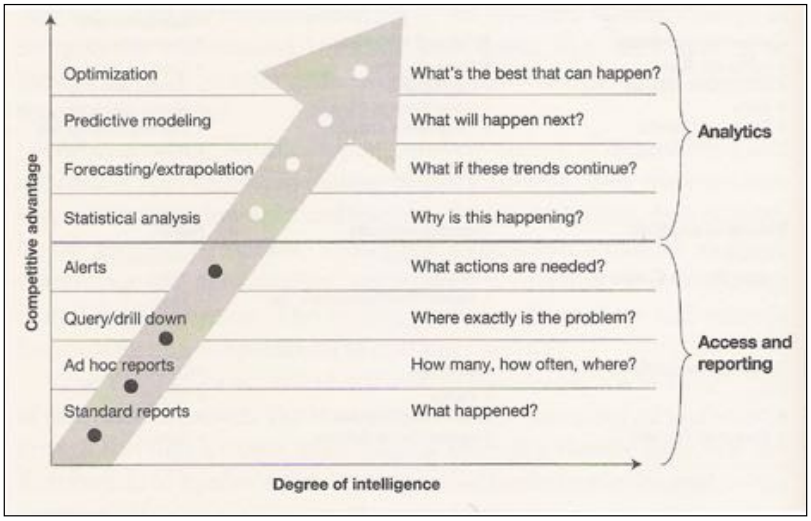
BUSINESS ANALYTICS

Big data is more than simply a matter of size: it is also about creating insights!

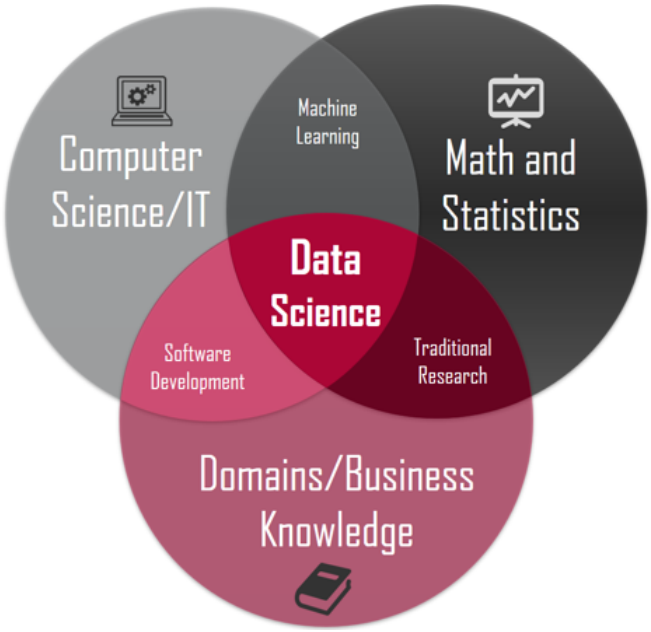


BUSINESS ANALYTICS

Big data is more than simply a matter of size: it is also about creating insights!



BUSINESS ANALYTICS



ANALYTICS: SHIFTING PARADIGMS

Do computers play more intelligently than humans?

Geplaatst op 23/10/2014 | Published in Schaken | Reageren

ANALYTICS: SHIFTING PARADIGMS

AlphaZero vs Stockfish / artwork by Chess.com.

Google's AlphaZero Destroys Stockfish In 100-Game Match

FM MikeKlein

Dec 6, 2017, 12:50 PM | 341 | Chess Event Coverage

English

Chess changed forever today. And maybe the rest of the world did, too.

ANALYTICS: SHIFTING PARADIGMS

Google AI algorithm masters ancient game of Go

Deep-learning software defeats human professional for first time.

Elizabeth Gibney

27 January 2016

ANALYTICS: SHIFTING PARADIGMS

Zondag 01 december 2019 | Het laatste nieuws het eerst op NU.nl

Professionele gospeler stopt carrière na nederlaag tegen Google-software

27 november 2019 15:28
Laatste update: 2 dagen, 22 uur geleden

De Zuid-Koreaanse gospeler Lee Se-dol stopt met zijn professionele carrière in het Chinese bordspel. De topsporter zegt tegen het Zuid-Koreaanse persbureau *Yonhap* dat zijn keuze vooral gemotiveerd is door zijn nederlaag tegen software van het Google-zusterbedrijf DeepMind in 2016.

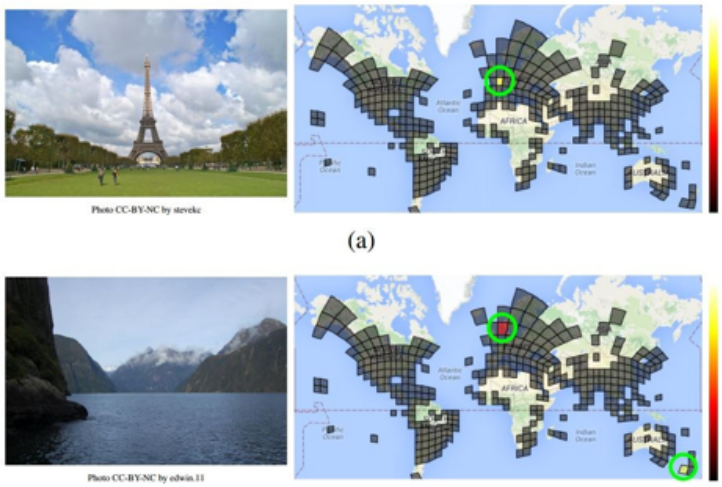
ANALYTICS: SHIFTING PARADIGMS

Computing

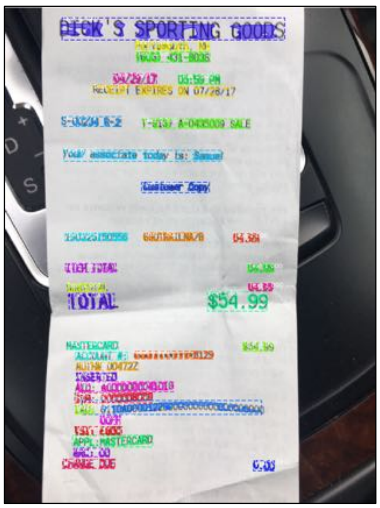
Google Unveils Neural Network with "Superhuman" Ability to Determine the Location of Almost Any Image

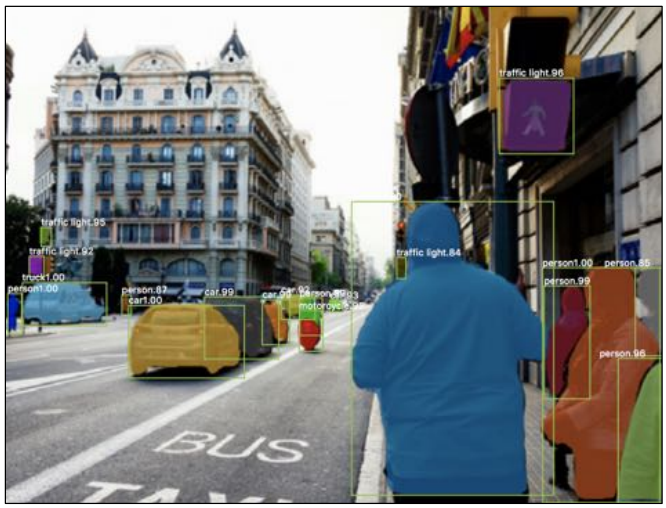
Guessing the location of a randomly chosen Street View image is hard, even for well-traveled humans. But Google's latest artificial-intelligence machine manages it with relative ease.

by Emerging Technology from the arXiv
February 24, 2016



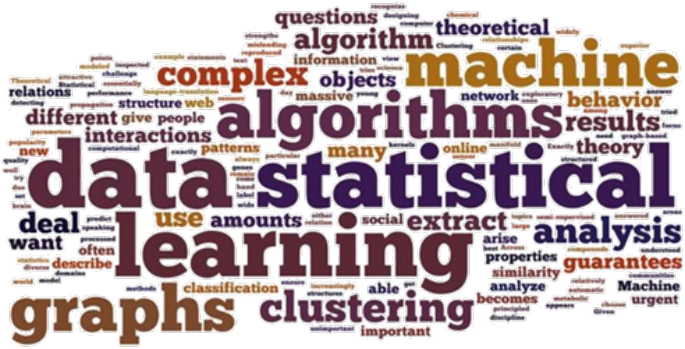
ANALYTICS: SHIFTING PARADIGMS



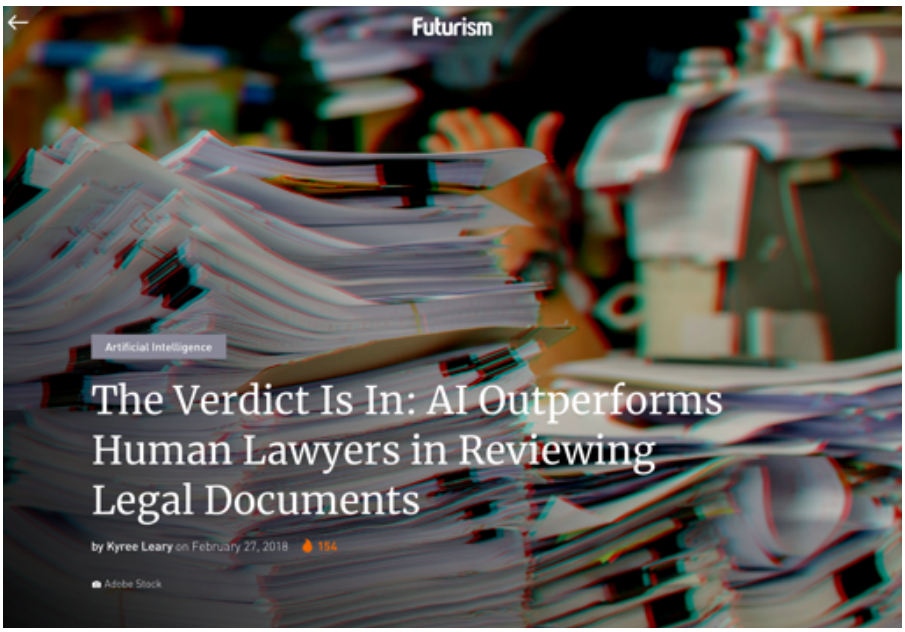


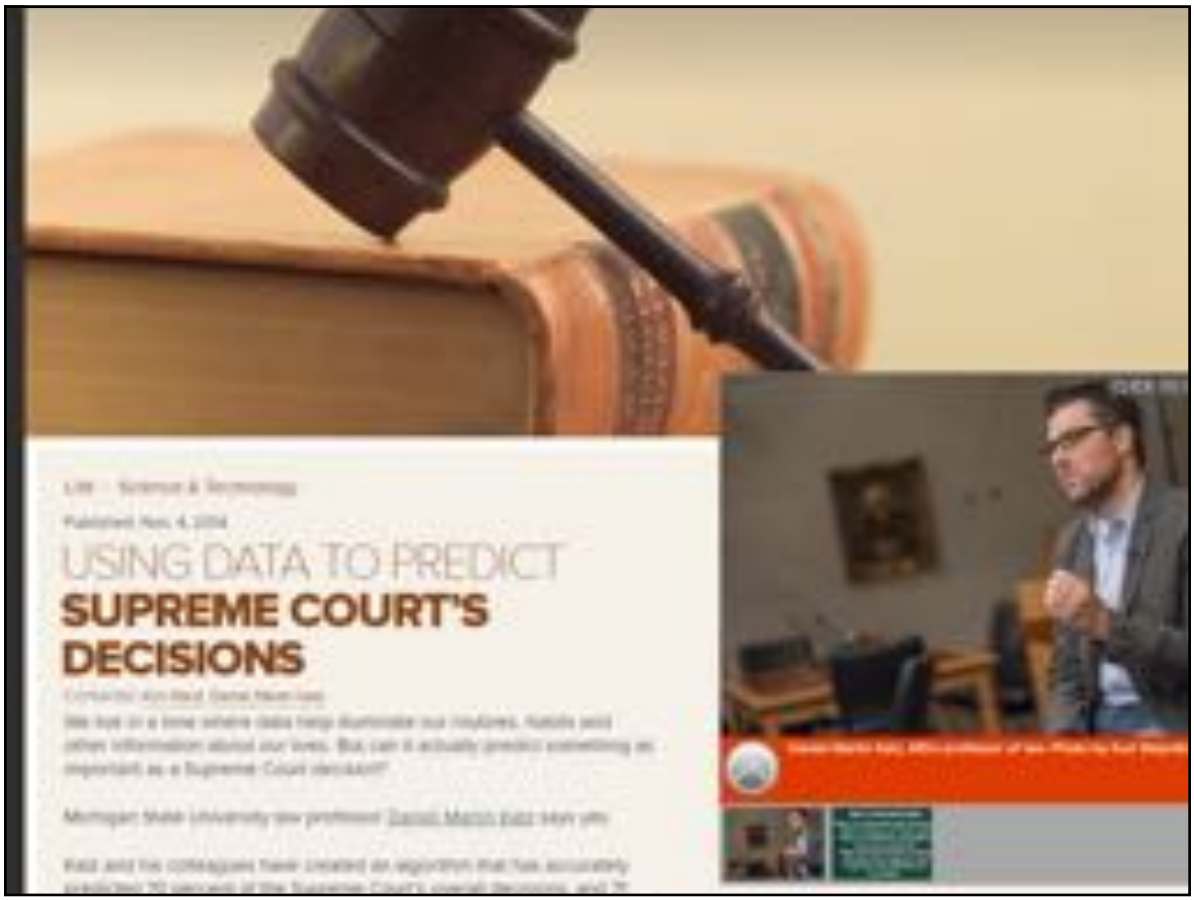
ANALYTICS: SHIFTING PARADIGMS

- Panama papers (2.6 TB, 11.5 billion documents)
  - > 5 billion emails
  - > 3 billion databases
  - > 2 billion PDF documents
  - > 1 billion images
  - > 320,166 text files
  - > 2,242 other files



ANALYTICS: SHIFTING PARADIGMS

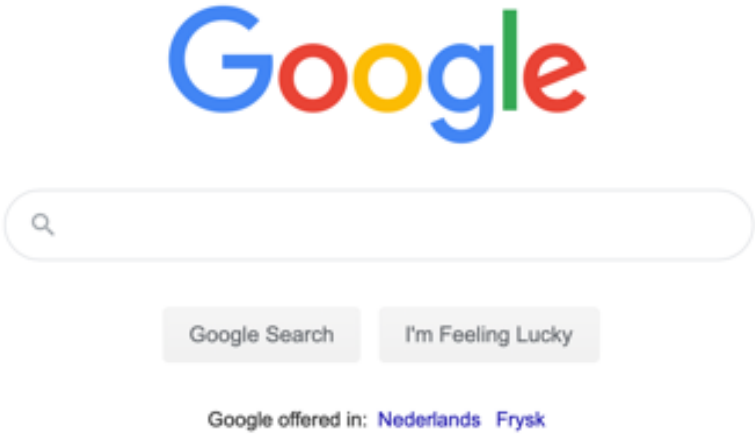




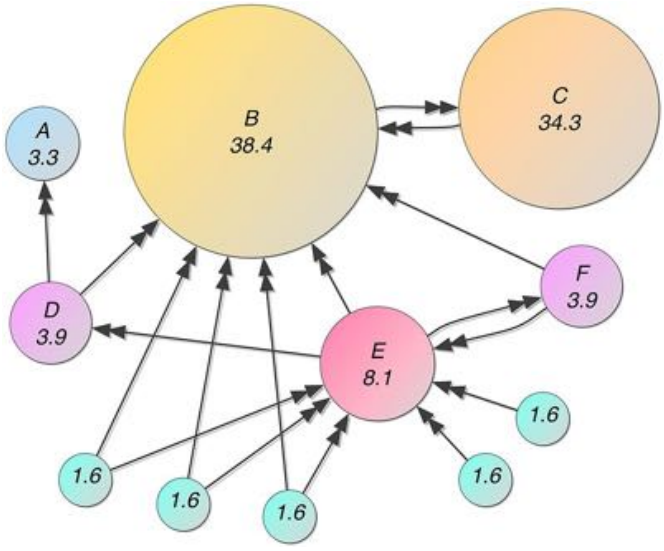
ANALYTICS: SHIFTING PARADIGMS

ANALYTICS: SHIFTING PARADIGMS

EXAMPLE 1: MARKOV MODELS



EXAMPLE 1: MARKOV MODELS



$$S(V_i) = (1-d) + d \sum_{j \in \text{In}(V_i)} \frac{1}{|\text{Out}(V_j)|} S(V_j)$$

EXAMPLE 1: MARKOV MODELS

- The same trick can be applied to sentences
- However, there are no links in sentences

$$\text{Similarity}(S_i, S_j) = \frac{|\{w_k | w_k \in S_i \text{ and } w_k \in S_j\}|}{\log(|S_i|) + \log(|S_j|)}$$

- This creates a system for summarization

EXAMPLE 1: TEXT SUMMARIZATION

BUSINESS NEWS FEBRUARY 20, 2020 / 12:09 PM / UPDATED 16 HOURS AGO

UBS's tech-savvy new boss Hamers not always loved by the Dutch

Bart H. Meijer, Michael Shields 4 MIN READ

AMSTERDAM/ZURICH (Reuters) - Ralph Hamers made his mark as a tech-savvy chief executive who led Dutch bank ING (INGA.AS) back to profitability after cementing its position as a no-frills lender in Germany and simplifying its product portfolio.

EXAMPLE 1: TEXT SUMMARIZATION

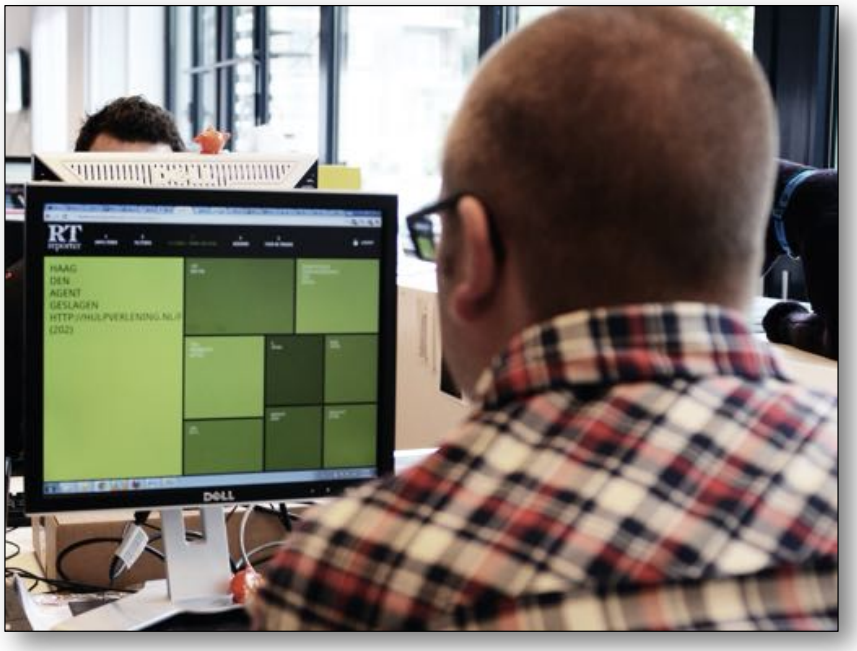
‘Black in Space’ Looks at Final Frontier of Civil Rights

In 1959, Ronald Erwin McNair walked into a South Carolina library. The 9-year-old aspiring astronaut wanted to check out a calculus book, but a librarian threatened to call the police if he didn't leave. McNair was black.

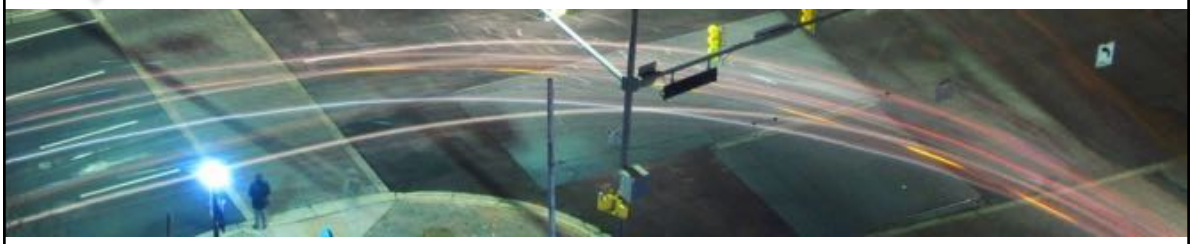
ASSOCIATED PRESS

PUBLISHED 20 FEBRUARY 2020

EXAMPLE 1: FROM NU.NL TO STRAKS.NL



EXAMPLE 1: TWITTER FOR TAXI

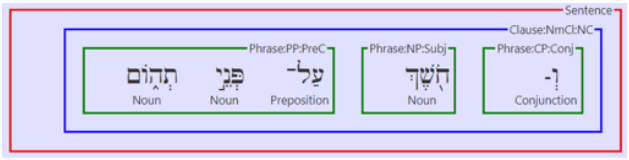


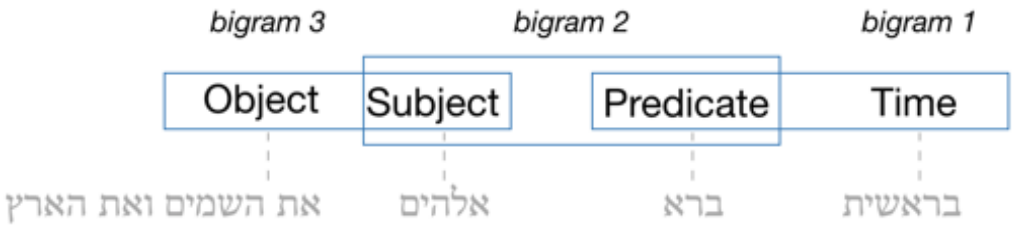
EXAMPLE 1: BIBLICAL HEBREW

EXAMPLE 1: BIBLICAL HEBREW

- ETCBC annotation data using Text-Fabric
- Word: part-of-speech: noun, article, preposition, …
- Phrase: function: object, subject, adjunct, …
- Phrase: type: verbal (VP), nominal (NP), …
- Clause: type: Way+X, Way+Qtl, …
- Clause: domain: Narrivate (N), Discursive (D), Quotation (Q), and Unknown (?)

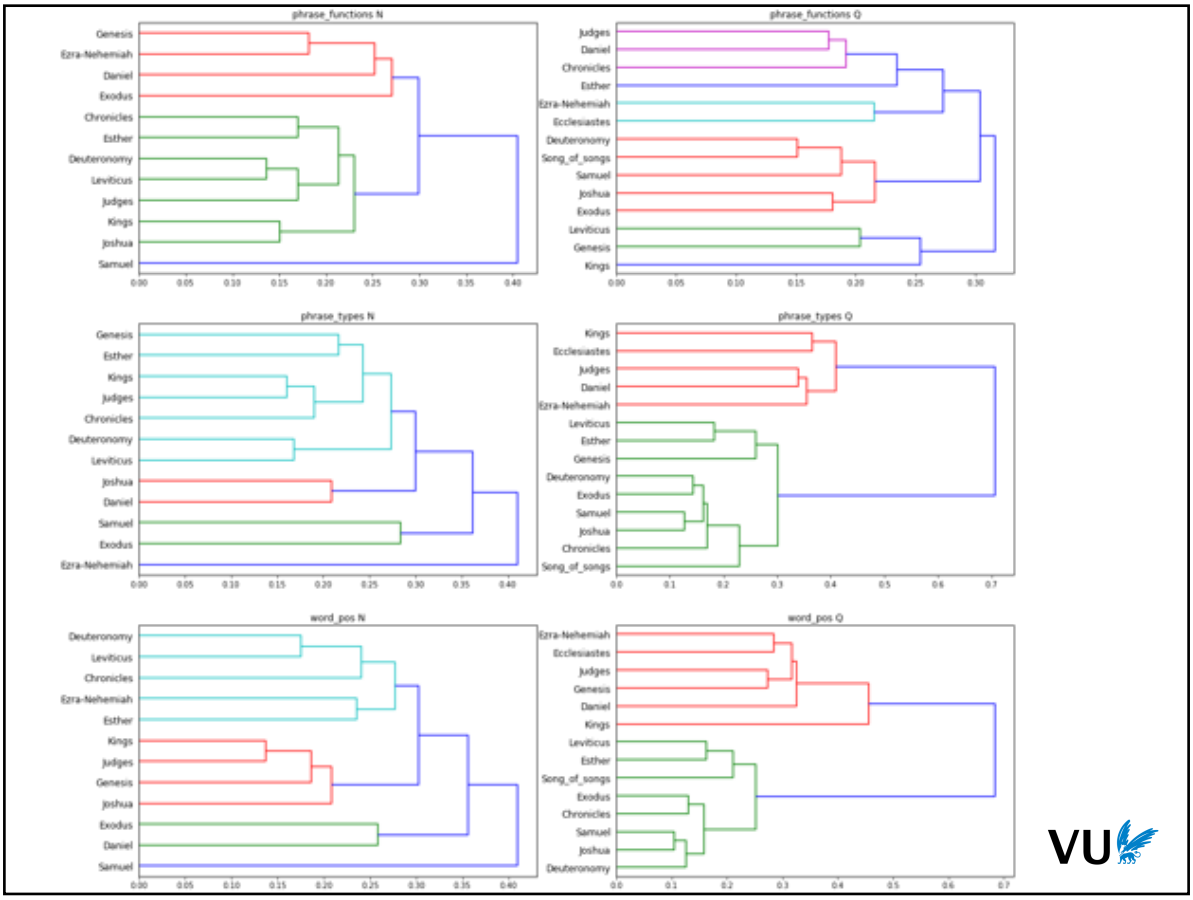
EXAMPLE 1: BIBLICAL HEBREW



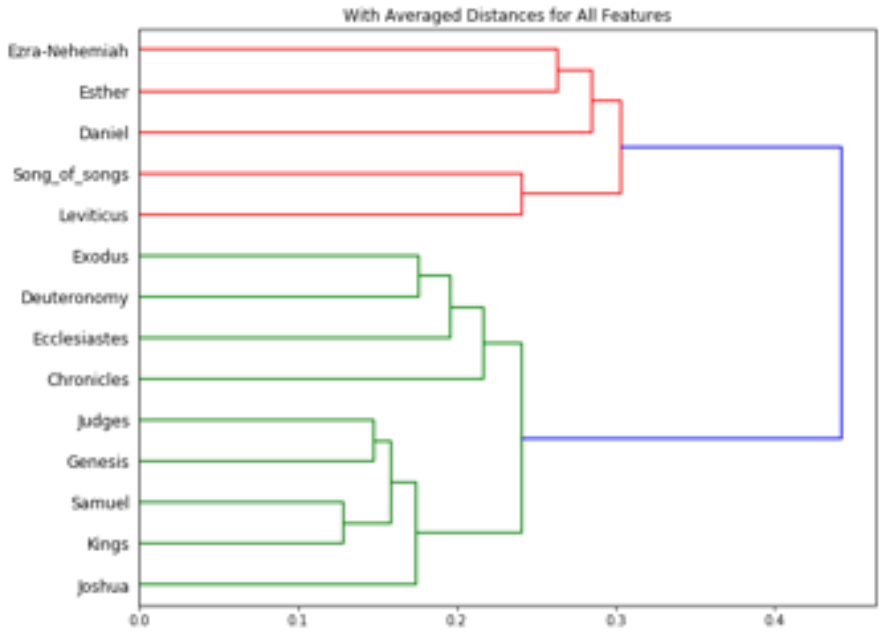


EXAMPLE 1: BIBLICAL HEBREW

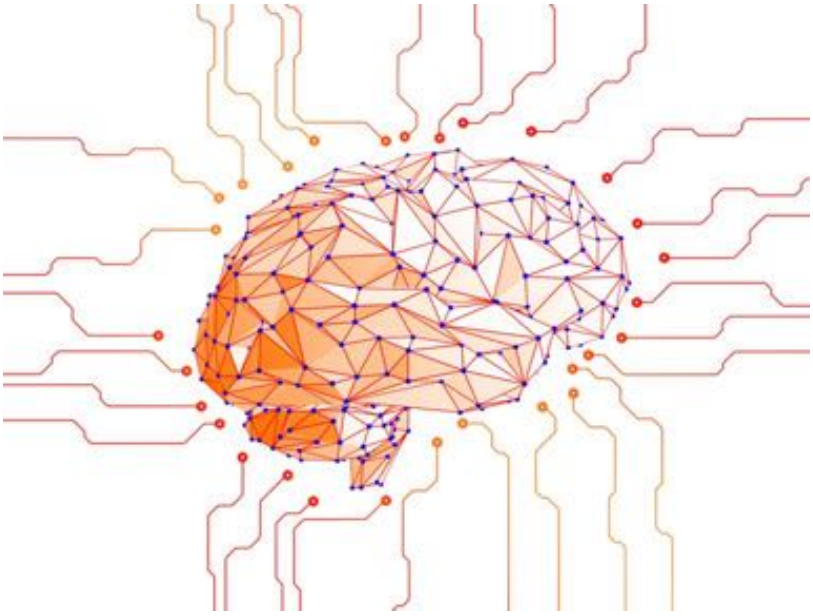
- Transition Matrices are constructed for the domains N and Q on part-of-speech and phrase function/type level



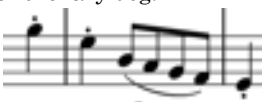
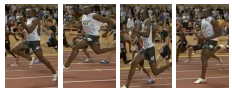
EXAMPLE 1: BIBLICAL HEBREW



EXAMPLE 2: NEURAL NETWORKS

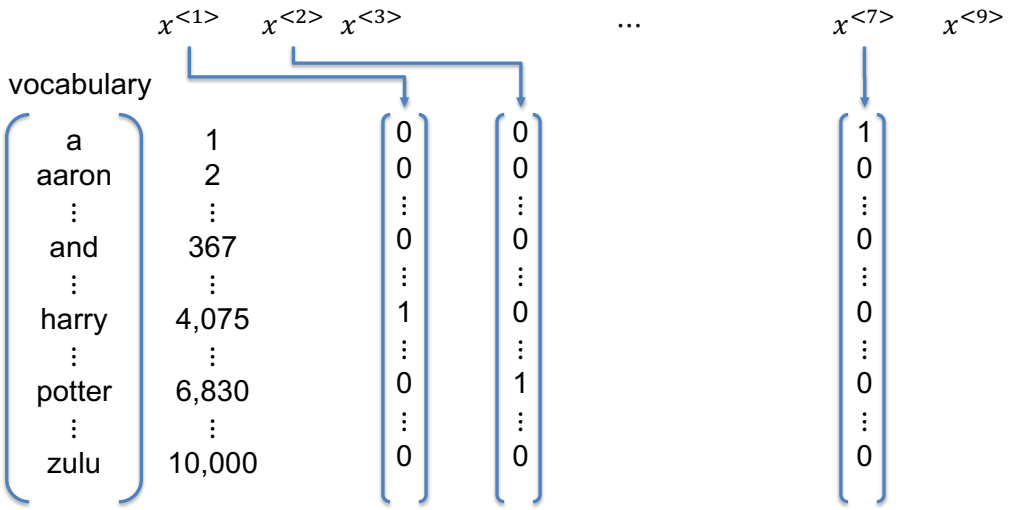


EXAMPLE 2: NEURAL NETWORKS

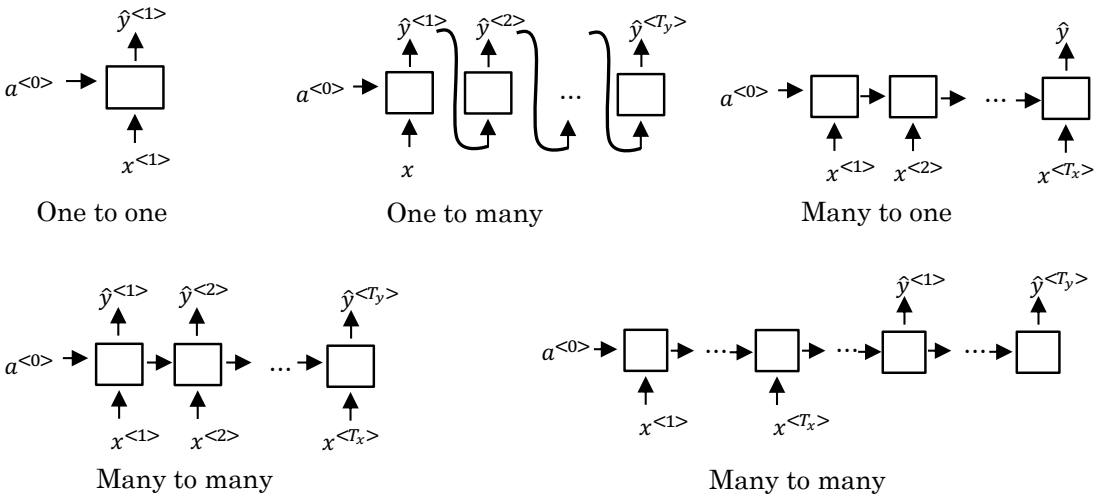
Speech recognition		→	"The quick brown fox jumped over the lazy dog."
Music generation	∅	→	
Sentiment classification	"There is nothing to like in this movie."	→	★☆☆☆☆
DNA sequence analysis	AGCCCCTGTGAGGAACTAG	→	AGCCCCTGTGAGGAACTAG
Machine translation	Voulez-vous chanter avec moi?	→	Do you want to sing with me?
Video activity recognition		→	Running
Name entity recognition	Yesterday, Harry Potter met Hermione Granger.	→	Yesterday, Harry Potter met Hermione Granger.

EXAMPLE 2: NEURAL NETWORKS

x: Harry Potter and Hermione Granger invented a new spell.



EXAMPLE 2: RECURRENT NEURAL NETWORKS



EXAMPLE 2: RECURRENT NEURAL NETWORKS

- Vocabulary: V = [a, aaron, …, zulu, <UNK>]

Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)

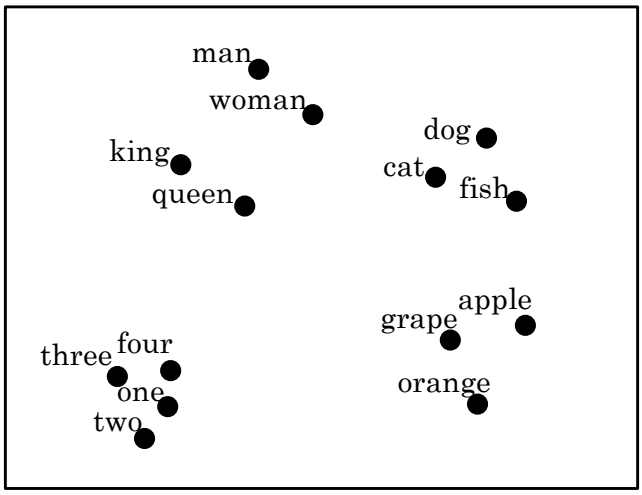
I want a glass of orange _____.

I want a glass of apple_____.

EXAMPLE 2: RECURRENT NEURAL NETWORKS

	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
Gender	-1	1	-0.95	0.97	0.00	0.01
Royal	0.01	0.02	0.93	0.95	-0.01	0.00
Age	0.03	0.02	0.7	0.69	0.03	-0.02
Food	0.09	0.01	0.02	0.01	0.95	0.97

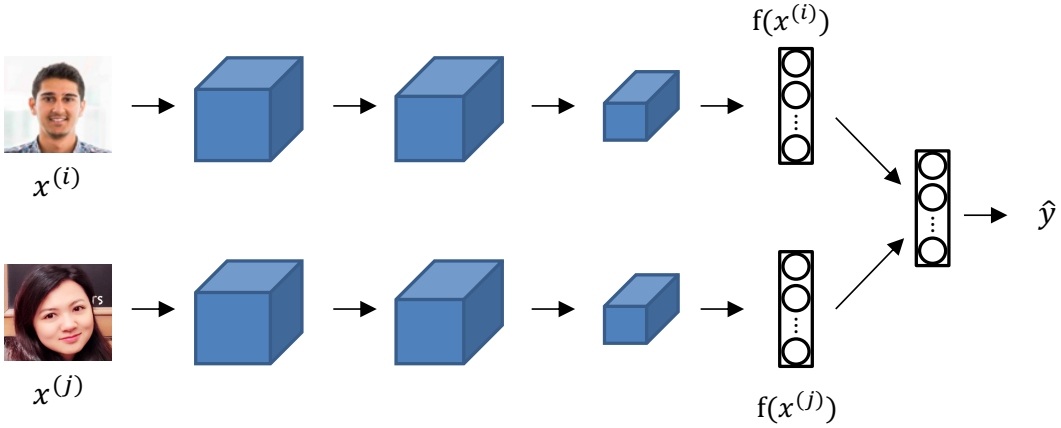
EXAMPLE 2: RECURRENT NEURAL NETWORKS



Visualization with t-SNE

EXAMPLE 2: RECURRENT NEURAL NETWORKS

- Relation to face encoding



VU

EXAMPLE 2: RECURRENT NEURAL NETWORKS

- Analogy to word embedding

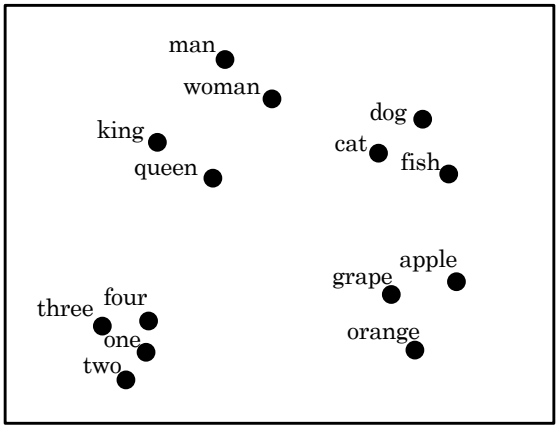
	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
Gender	−1	1	-0.95	0.97	0.00	0.01
Royal	0.01	0.02	0.93	0.95	-0.01	0.00
Age	0.03	0.02	0.70	0.69	0.03	-0.02
Food	0.09	0.01	0.02	0.01	0.95	0.97

- man → woman as king → ??

VU

EXAMPLE 2: RECURRENT NEURAL NETWORKS

- Analogy to word embedding



$$e_{man} - e_{woman} \approx e_{king} - e_?$$

EXAMPLE 2: RECURRENT NEURAL NETWORKS

- Similarity function: $sim(u, v) = \frac{u^T v}{\|u\|_2 \ \|v\|_2}$

- Examples:
  - Man:Woman as Boy:Girl
  - Ottawa:Canada as Nairobi:Kenya
  - Big:Bigger as Tall:Taller
  - Yen:Japan as Ruble:Russia

EXAMPLE 2: BIBLICAL HEBREW

- ETCBC annotation data using Text-Fabric
- Word: part-of-speech: noun, article, preposition, …
- Word: verbal stem: Hif'il, Pi'el, Qal, …
- Phrase: function: object, subject, relative, …
- Clause: type: nominal, participle, wayyiqtol, …

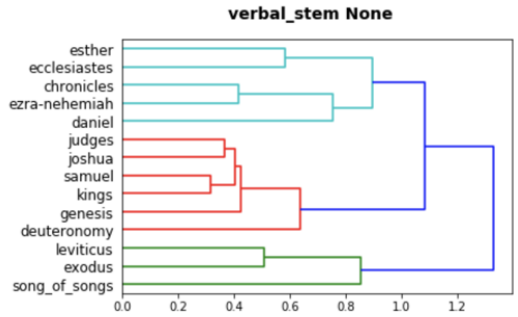
EXAMPLE 2: BIBLICAL HEBREW

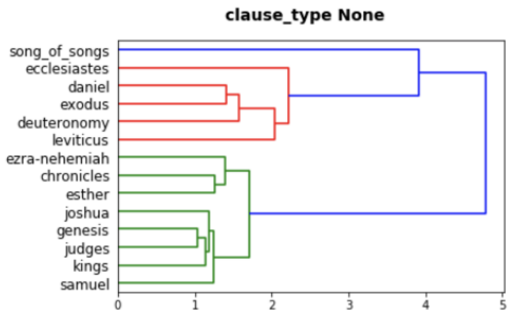
- LSTM model is trained on corpus:
  - > Genesis, Exodus, Deuteronomy, Leviticus, Judges, Joshua, Kings, Samuel
  - > Chronicles, Daniel, Ecclesiastes, Esther, Ezra-Nehemiah, Song of Songs

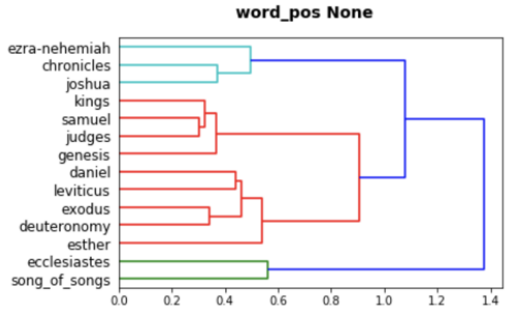
וַ יַּעַשׂ מֹשֶׁה ׃ כְּ כֹל אֲשֶׁר צִוָּה יְהוָה אֹתֹו כֵּן עָשָׂה

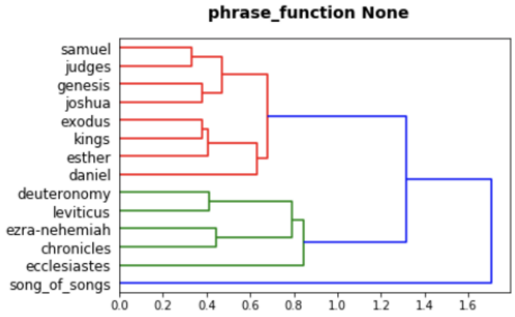
וּ בָאוּ עָלֶיךָ כָּל־ הַ בְּרָכֹות הָ אֵלֶּה וְ הִשִּׂיגֻךָ כִּי תִשְׁמַע בְּ קֹול יְהוָה אֱלֹהֶיךָ

EXAMPLE 2: BIBLICAL HEBREW

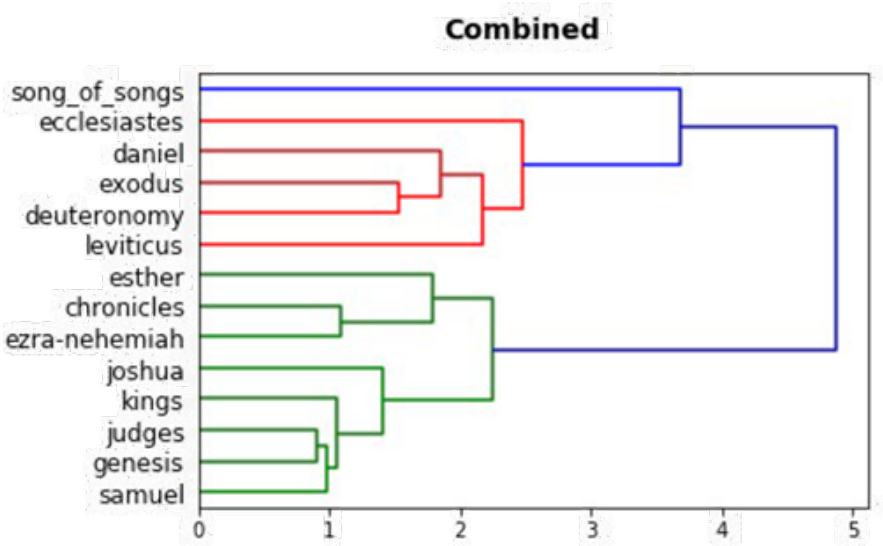








EXAMPLE 2: BIBLICAL HEBREW



QUESTIONS