

Algoritmer och maskininlärning

Olof Mogren

Chalmers tekniska högskola

2017

YAHOO!

Google

twitter

amazon.com®

facebook

De här företagen vill

- Tjäna pengar
- Hitta mönster i stora datamängder
- Göra förutsägelser
- Klassificera data
- Förstå människan

Maskininlärning

- Tjäna pengar
- Hitta mönster i stora datamängder
- Göra förutsägelser
- Klassificera data
- Förstå människan

Tjäna pengar (minikurs i ekonomi)

- Förstå vad dina kunder vill ha
- Ge dina kunder vad de vill ha
- Ta betalt för det de vill ha

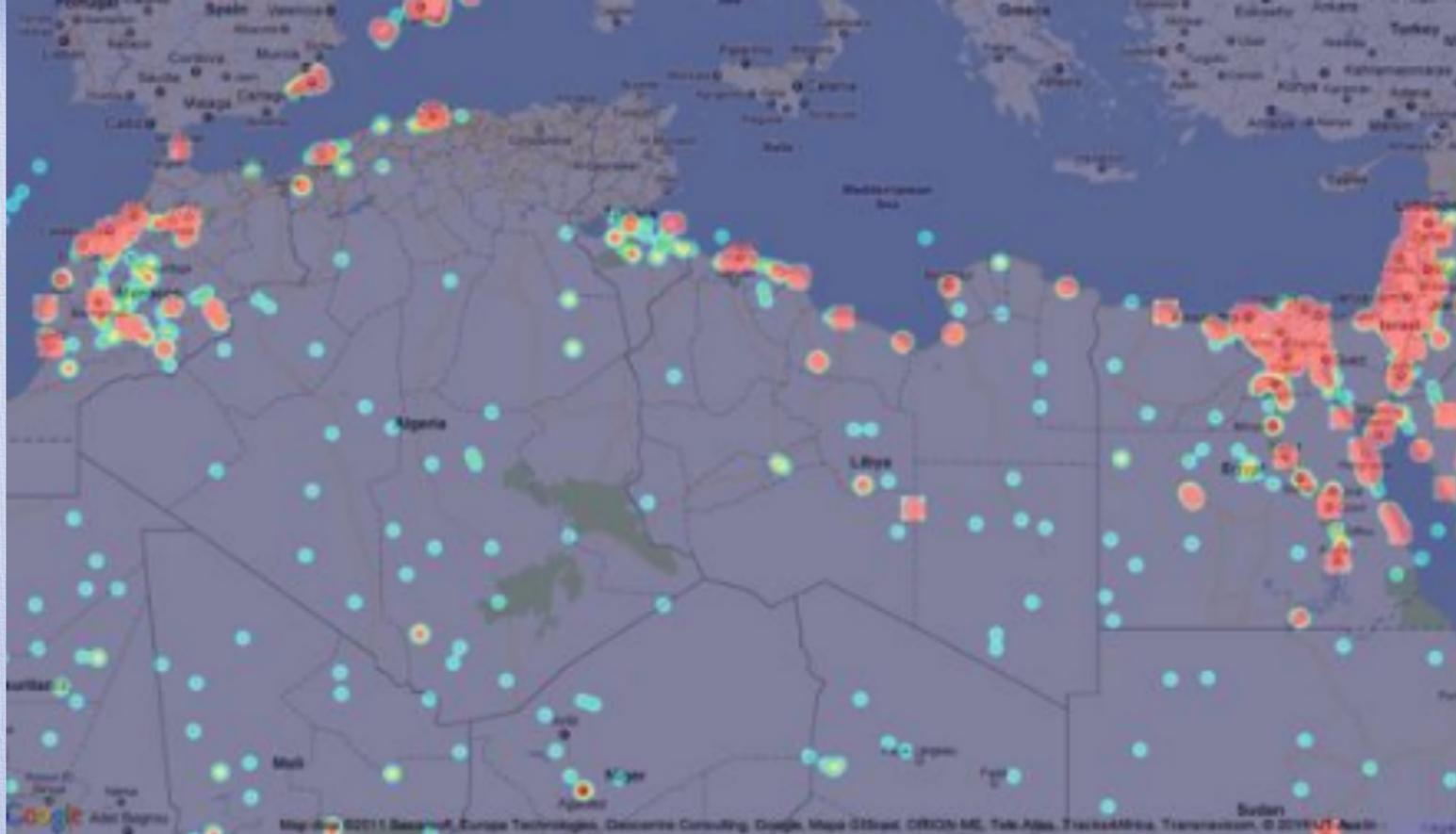


Hitta mönster i stora datamängder

- Twitter
- 58 miljoner tweets om dagen
- 671 tweets per sekund

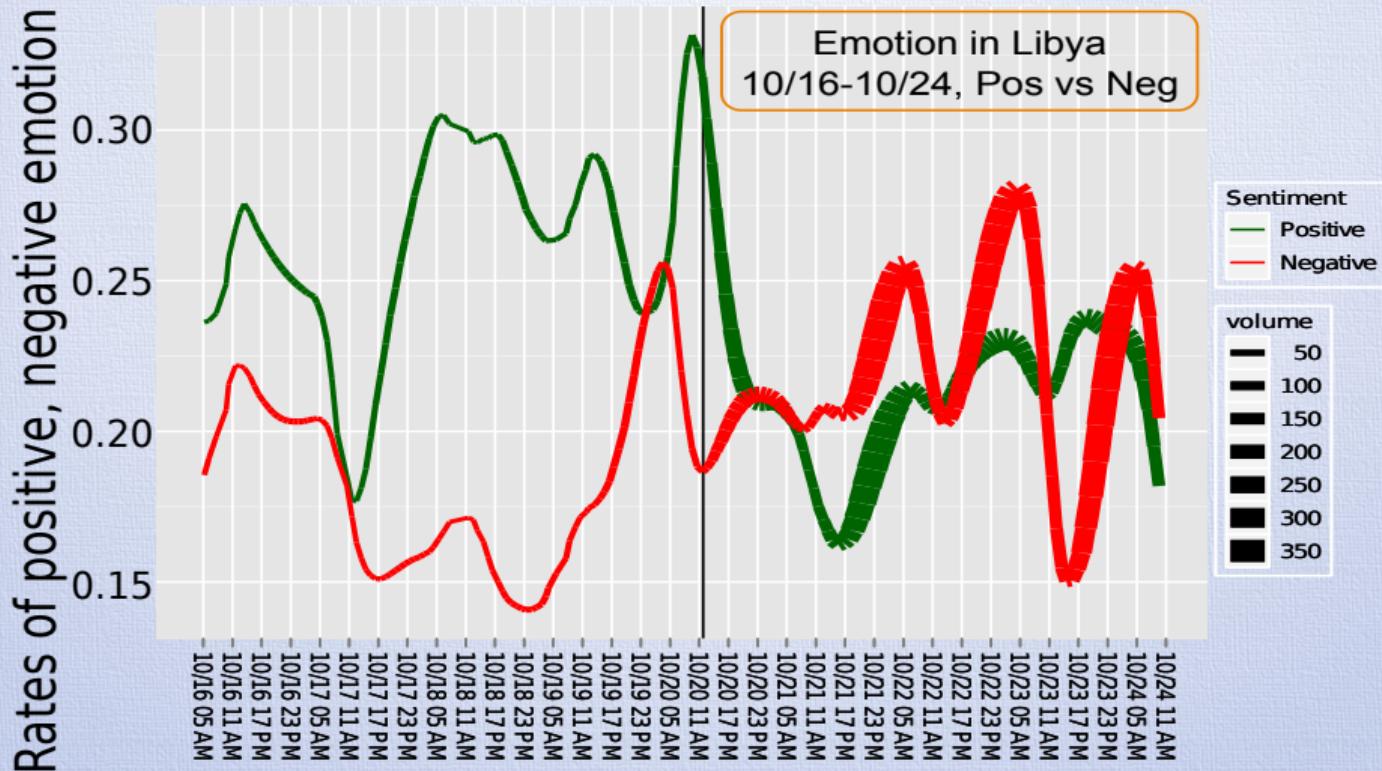
Information i tweets

- Ett tweet - lite information
- 58 miljoner tweets - mycket information
- Hur hantera?



Tweets under arabiska våren.
David Beaver, University of Austin, Texas.

<http://mogren.one/>



Beroenden

- Synliga observationer
- Dolda variabler
- Statistisk analys
- The Curly Fries Conundrum
(Jennifer Golbeck, TED Talk)



Hittills

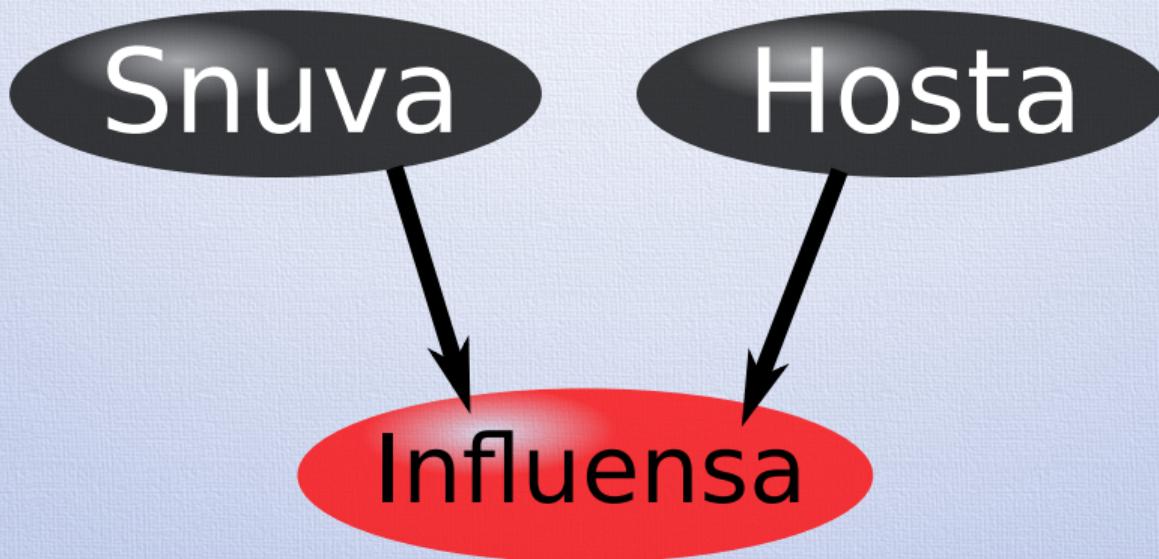
- Twitter, Google, övriga har massor av information
- Personifiering, riktad reklam
- Rekommendationer
- Företagsplanering (vad är coolt? inte coolt?)
- Maskininlärning: centralt

Maskininlärning



Maskininlärning

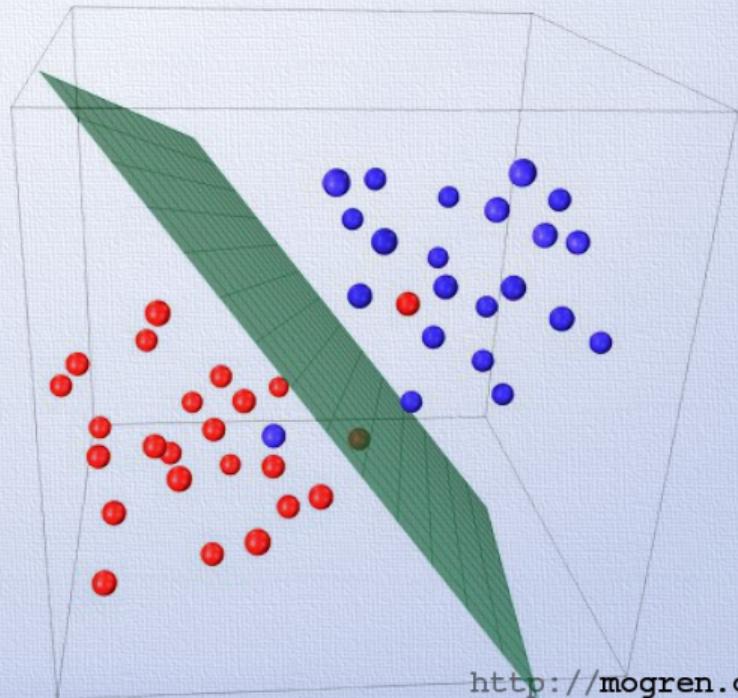
- Förbättrar prestanda efter observationer.
- Exempel: symptom, diagnoser



maskininlärning

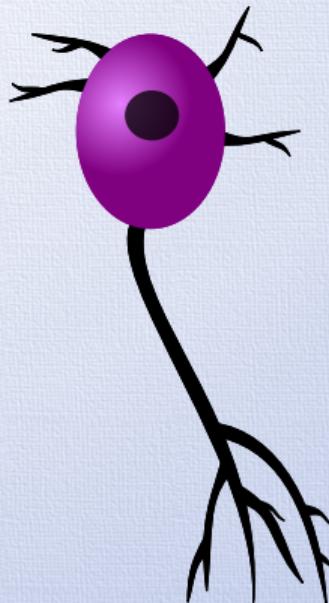
Typer:

- Övervakad (supervised)
 - Ex. klassificering
 - Ex. förutsägelser
- Oövervakad (unsupervised)
 - Ex. klustering
 - Ex. word2vec - ordvektorer



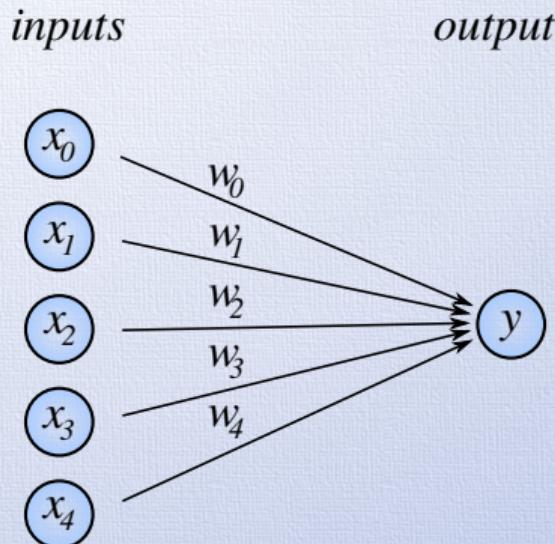
Artificiella neurala nätverk

- Maskininlärning inspirerad av hjärnan
- Artificiella neuroner
- Signaler



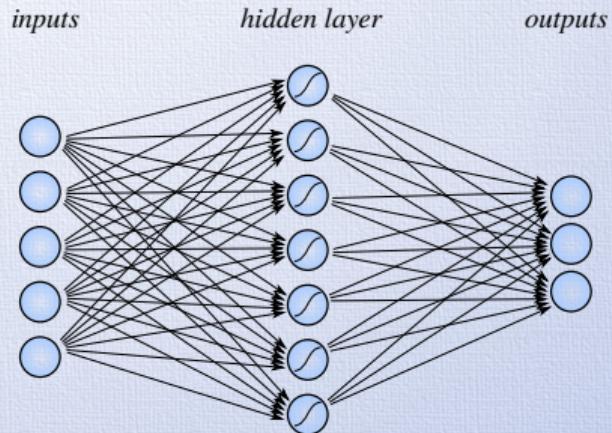
Artificiella neurala nätverk

- Numeriska signaler
- Neuronerna lär sig



Deep Learning

- Neuroner i lager
- Många lager → djupt
nätverk
- Universal approximation
- Feature learning



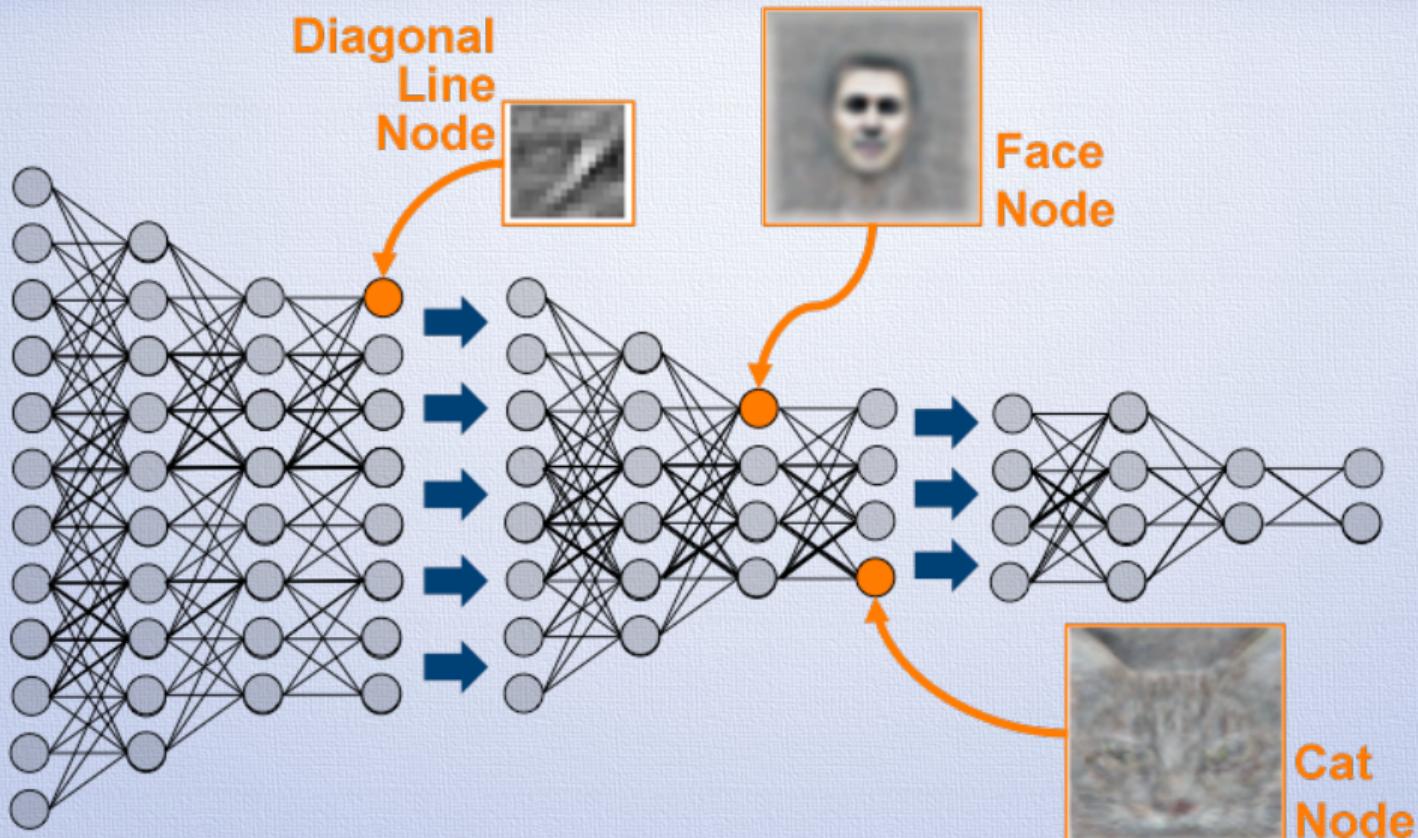


Bild från författarna.

<http://mogren.one/>

Deep Learning

- Översättningssystem
- Text till tal
- Tal till text
- Genererad handstil
 - wikimarkup
 - programmeringskod
- Bildklassificering - ImageNet
 - bättre än människan
- Gmail "smart reply"

Deep learning i javascript



cs231n.stanford.edu



playground.tensorflow.org

Neural assistent för diskussionsforum

(Hagstedt, Mogren, 2016)

- Djupa neurala sekvensmodeller
- Numeriska vektorvärda representationer
- Rekommendera relaterade
 - Inlägg
 - Trådar
 - Användare
- Nystartat företag: Textual

2:33 PM
Anyone with experience using tensorflow when creating neural networks?

BOT 2:33 PM
Seems like [REDACTED] (@[REDACTED]) and [REDACTED] (@[REDACTED]) have asked a similar question! Maybe ask one of them? 😊

Was this information helpful? Please give it a vote!

1
Seems like [REDACTED] over at #_[REDACTED]_machinelearning and #_[REDACTED]_data have asked a similar question! Maybe ask over there? 🚶

1
I found a couple of similar posts to your question! 🌐

Posted in	Posted in
_[REDACTED]_machinelearning	_[REDACTED]_machinelearning
by [REDACTED]	by [REDACTED]
Anyone checked out	word2vec?
TensorFlow from	https://[REDACTED]
Google?	https://[REDACTED]
https://[REDACTED]	https://[REDACTED]
https://[REDACTED]	https://[REDACTED]
https://[REDACTED]	https://[REDACTED]

1

Atari med Deep Q Learning

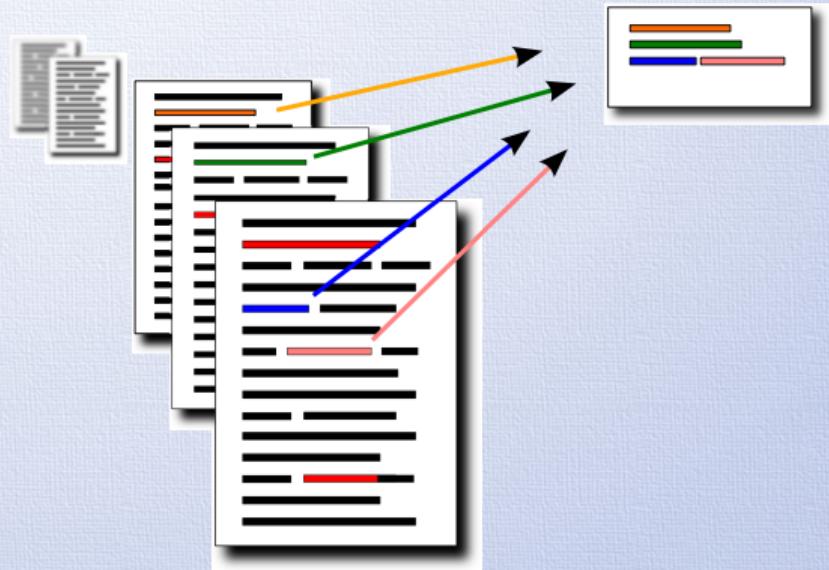


Online Offline

<http://mogren.one/>

Multi-Document Summarization

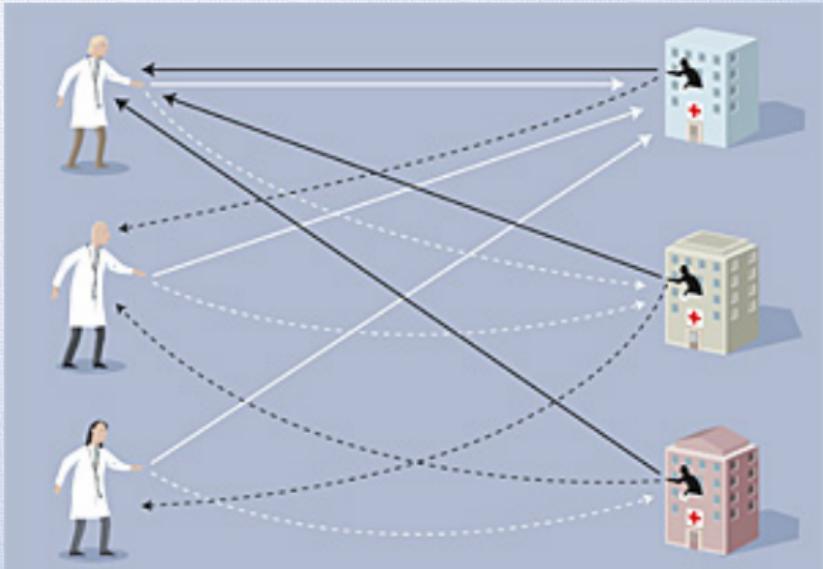
- Extrahera viktiga meningar
- "Viktiga" meningar - hög likhet med indata
- Neural word embeddings



Algoritmer

The stable marriage problem

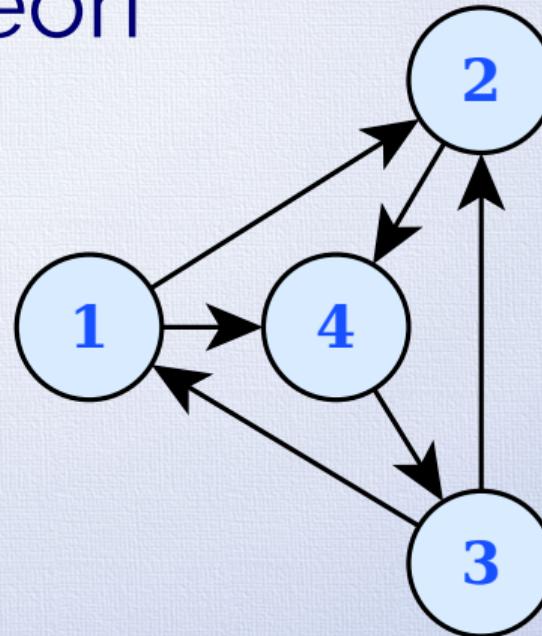
- Gale-Shapley 1962
- Ekonomipriset till Nobels minne



Johan Jarnestad
Kungl. Vetenskapsakademien

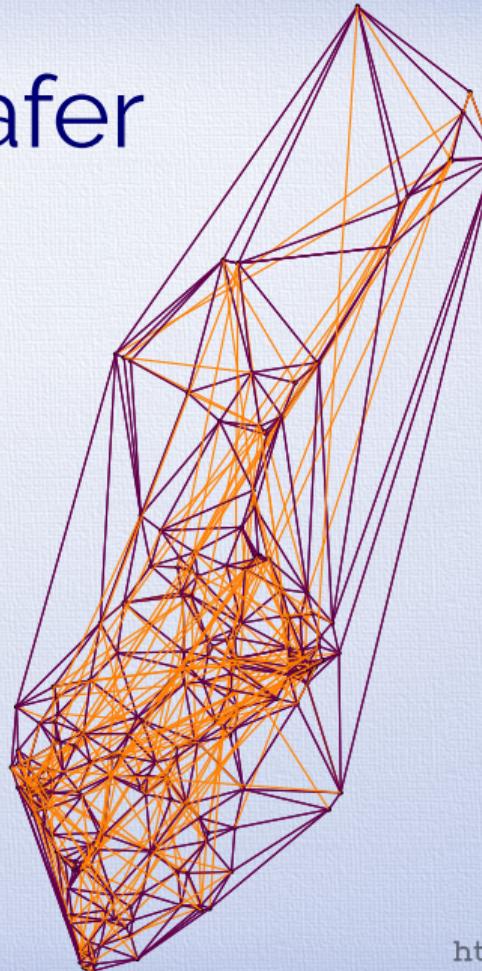
Grafteori

- Modeller för data
- Noder
- Kanter



Använda grafer

- Befolkningsdata SCB
- Proportionerlig täthet
- Simuleringar



Bioinformatik

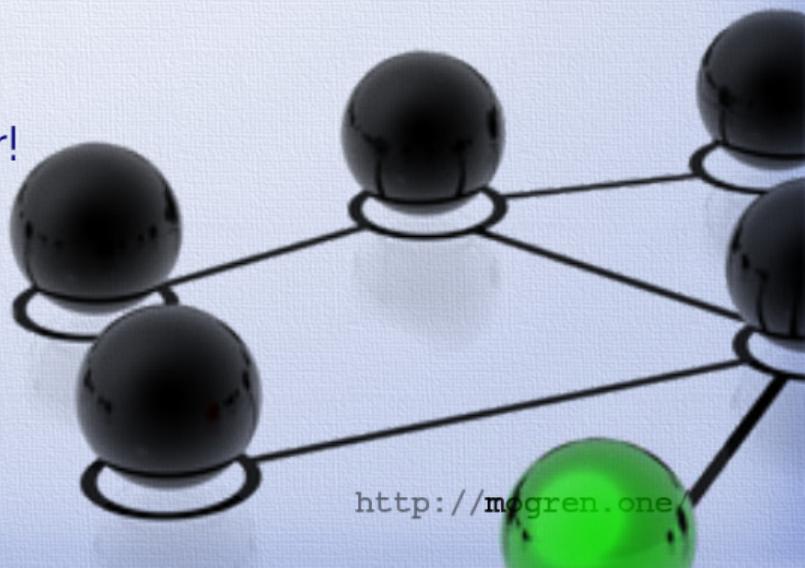
- DNA-sekvenser
- Bayesiska metoder
- Grafiska modeller (e.g.
HMMs)



<http://mogren.one/>

CSALL Masterprogram

- Algorithms, Languages and Logic
- Ett skitbra program, där allt kul sker!
- Kärnan av datavetenskap

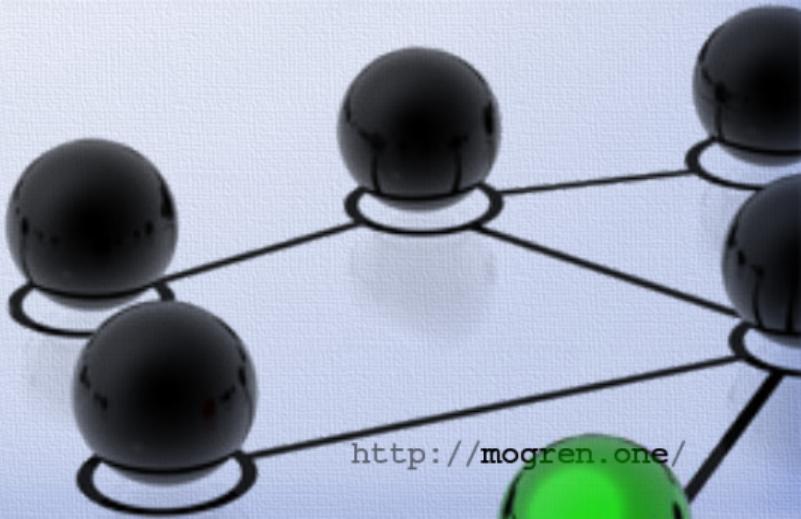


<http://mogren.one>

Algoritmer



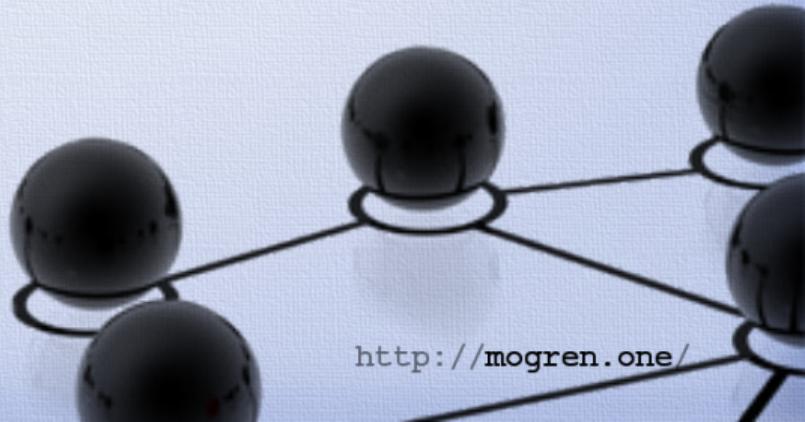
- Grundkurs, TIN093, DIT600
- Fortsättningskurs,
TDA251, DIT280
- Beräkningsbarhet
- Och annars?
- Peter Damaschke,
Muhammad Azam Sheikh



Maskininlärning



- TDA231, DIT380
- Probabilistiska grafiska modeller
- Support vector machines
- Deep learning
- Devdatt Dubhashi

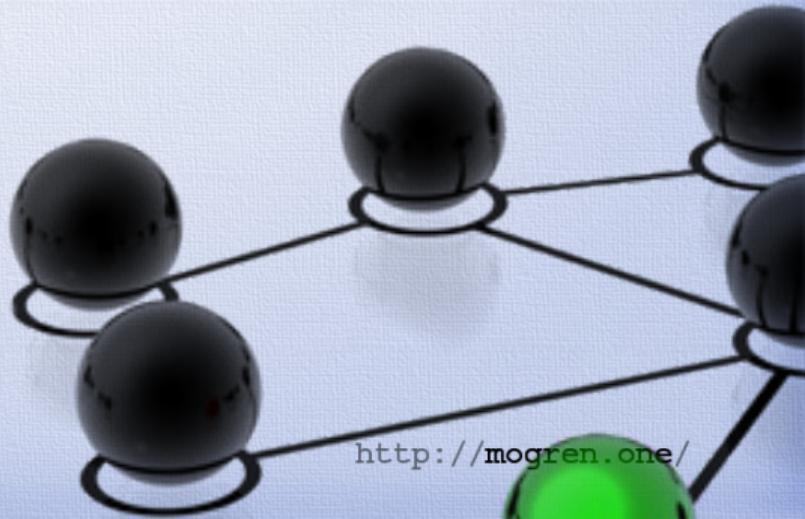


<http://mogren.one/>

Artificiell intelligens



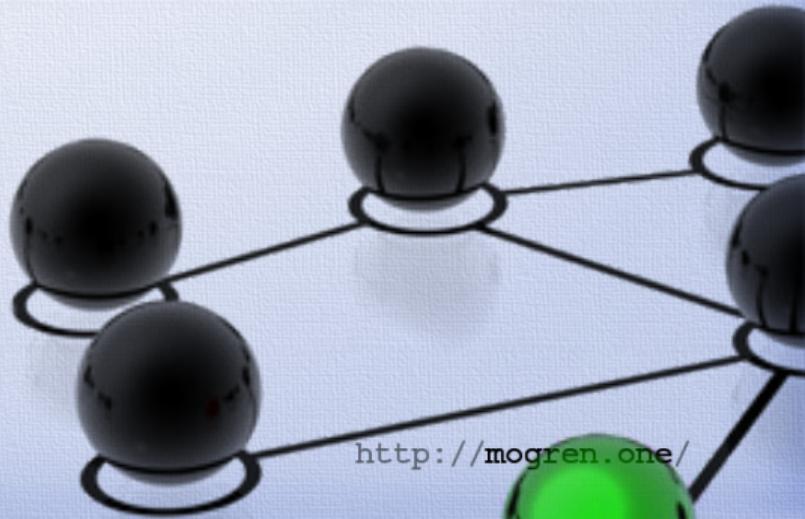
- TIN172,DIT410
- Teknisk kurs
- Programmeringsprojekt, Shrdlite
- Essä
- Peter Ljunglöf



Diskret optimering



- TDA206/DIT370
- Linear programming
- Essä
- Ashkan Panahi, Peter Damaschke



Neurala nätverk

- FFR135
- På institutionen för teknisk fysik
- "Ganska teoretisk"
- "Väldigt givande"
- Bernhard Mehlig





Reflektioner

- Plan: människa-dator-interaktion
- Upptäckte algoritmer
- Nu doktorand

Slutsatser

- Börja brett
- Ni hittar det som är kul
- Mattekurserna är värdefulla
- ...och kul!

Tack för er uppmärksamhet!



olof@mogren.one

<http://mogren.one/>

<http://www.cse.chalmers.se/research/lab/>

<http://mogren.one/>