

# Søketek- supportLearning!!!

2022-11-02

Classification  
Learning to rank  
Websøk 101 - link analysis  
Jobbe med oblig E





# Science fair

- Tema må submittes i dag
- Science fair om 2 uker
- På Python 10:15-14



# Oblig D-feedback

- Tar uvanlig lang tid
- Kommer (soon(tm))

# Support vector machines (SVMs)

- Ligger an til å ikke bli tilstrekkelig dekt i gruppetimene
- Relativt viktig for eksamen
- Oppfordrer til å fokusere effort hit

# Binary classifier

- En classifier med to klasser
- Bullet 2
- Bullet 3

# Feature-engineering og domenekunnskap

- Feature engineering: finn gode features(trekk) å lære på
- Domenekunnskap: Alt du a posteriori vet om noe
- **La clé**: utnytt domenekunnskap for å lage bedre features
- Eks: Skal man lære hvor rikt et land er gir det mening å se penger

# Websøk, link analysis

- Anta:
- Gode sider lenker til andre gode sider
- Dårlige sider peker til andre dårlige sider



# Internett som en graf

- Hver nettside er en node
- Hver link er en (rettet) kant
- Bullet 3



# PageRank

- En metode å ranke sider
- La en silver surfer surfe webgrafen random en vilkårlig mengde tid
- Jo mer tid silver surferen tilbragte på en side, jo høyere rankes siden.
- Mulig å beregne PageRank i  $O(1)$  med fancy matte(Utenfor pensum)





# Oblig E-arbeid

Frist neste fredag.

Siste oblig! 😊

