

Bruk av grafbasert maskinlæring til deteksjon av hvitvasking

NR — Fagseminar om datadrevet
antihvitvasking og svindeldeteksjon

Fredrik Johannessen

Tema: Vår nylig publiserte artikkel



Contents lists available at [ScienceDirect](#)

The Journal of Finance and Data Science

journal homepage: www.keaipublishing.com/en/journals/jfds



Research Article

Finding money launderers using heterogeneous graph neural networks



Fredrik Johannessen^a, Martin Jullum^{b,*}

^a DNB, P.O. Box 1600, Sentrum, NO-0021, Oslo, Norway

^b Norwegian Computing Center, P.O. Box 114, Blindern, NO-0314, Oslo, Norway

Lenke: <https://doi.org/10.1016/j.jfds.2025.100175>

Bakgrunn

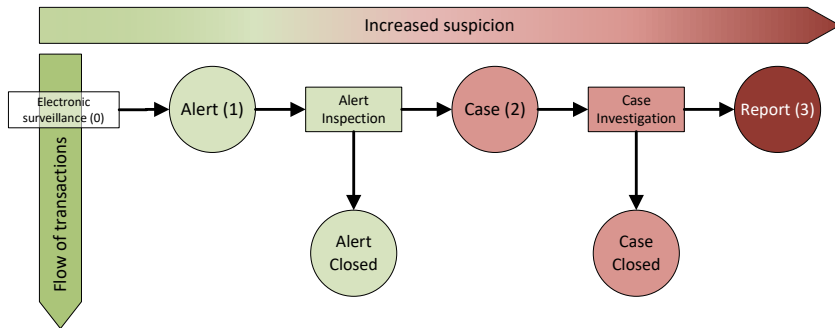
Hvitvasking

- Å gjøre utbytte fra kriminelle handlinger tilsynelatende lovlig ved å skjule dets opprinnelse
- Et omfattende globalt problem: $\sim 2\text{--}5\%$ av globalt BNP
- Anti-money laundering (AML) lover: Tiltak som banker må iverksette for å hindre at de blir utnyttet til hvitvasking



Elektronisk overvåking av transaksjoner

- Mål: Å avdekke transaksjoner tilknyttet hvitvasking



Regelbasert elektronisk overvåking

- Et sett med regler som består av noen få if/else-betingelser med faste terskelverdier som avgjør om et varsel utløses
- Utgjør fortsatt en betydelig del av de elektroniske overvåkingssystemene i banker

Example

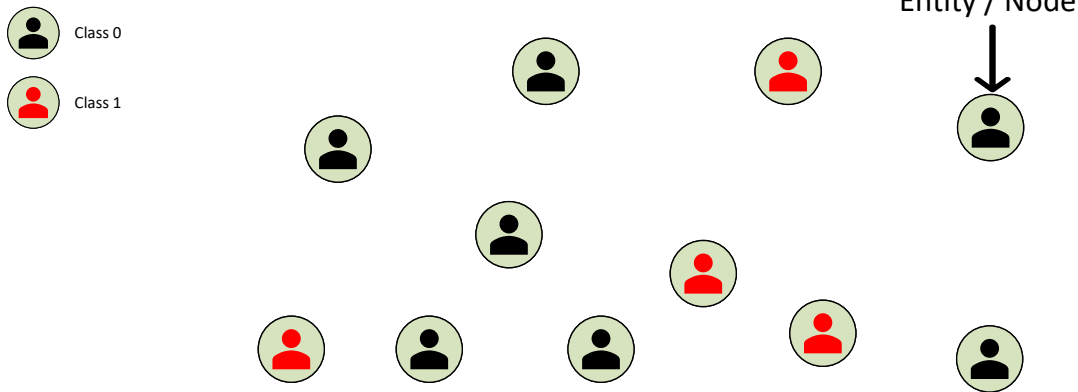
Hvis

- transaksjonen overstiger 50 000 kr, og
- den går til et land på listen over skatteparadiser, og
- kunden som utfører transaksjonen er klassifisert som privatkunde, så utløses et varsel.

Maskinlæring & nettverksvitenskap

Maskinlæring & nettverksvitenskap

- Maskinlæringsalgoritmer (ML) som automatisk lærer når varsler skal utløses
- Tema for denne presentasjonen: veiledet læring



Maskinlæring & nettverksvitenskap

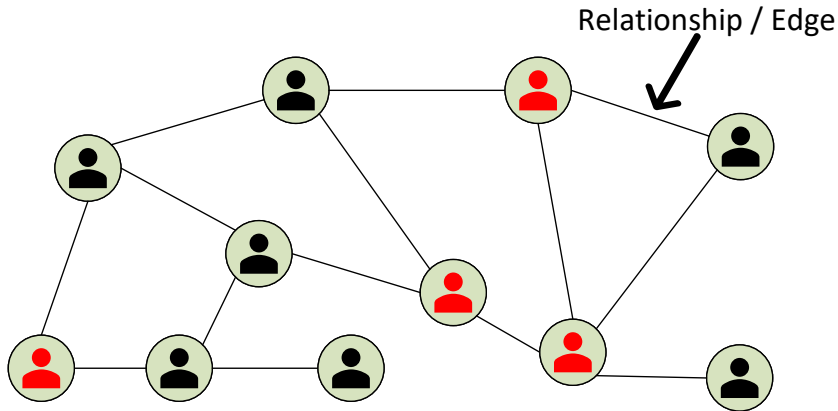
- Utnyttelse av relasjoner mellom entiteter
- Nettverk $\approx$ graf



Class 0



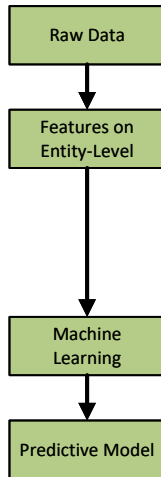
Class 1



Maskinlæring & nettverksvitenskap

ML på entitetsnivå

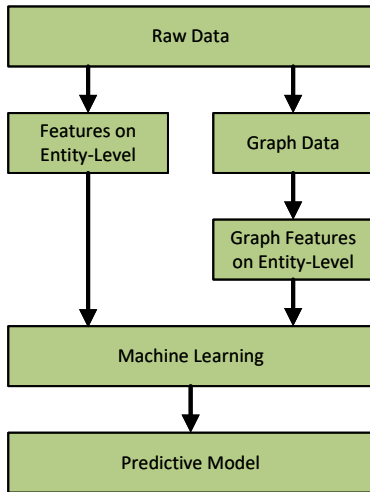
- Generere variabler på entitetsnivå
- Variablene brukes i en prediktiv modell



Maskinlæring & nettverksvitenskap

ML med tilleggsvariabler fra grafen

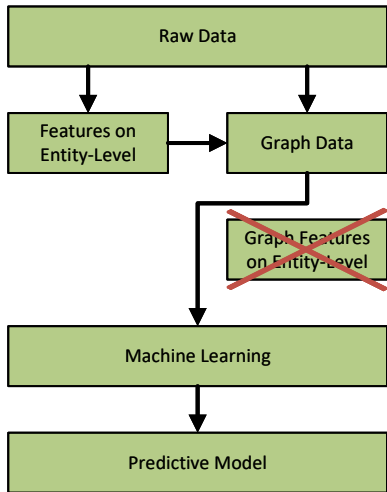
- Bruke grafen til å lage nye variabler
- Variablene fanger entitetens rolle i grafen



Maskinlæring & nettverksvitenskap

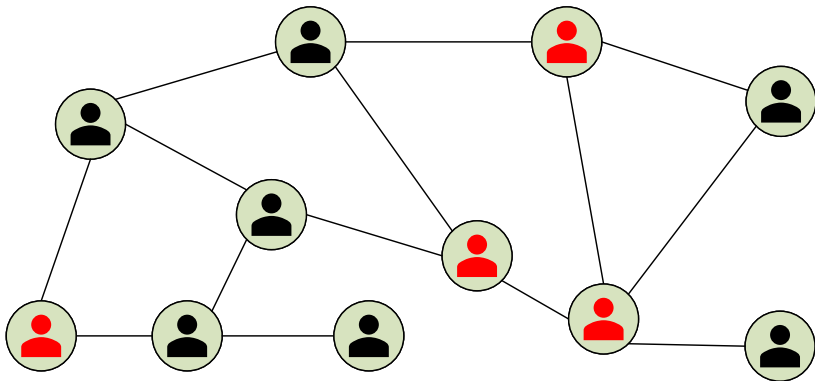
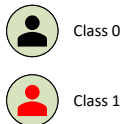
Maskinlæring direkte på grafen

- Trene en ML-modell direkte på grafen
- Graph Neural Network (GNN)
 - ✓ State-of-the-art
 - ✓ Kombinerer graf- og entitetsdata

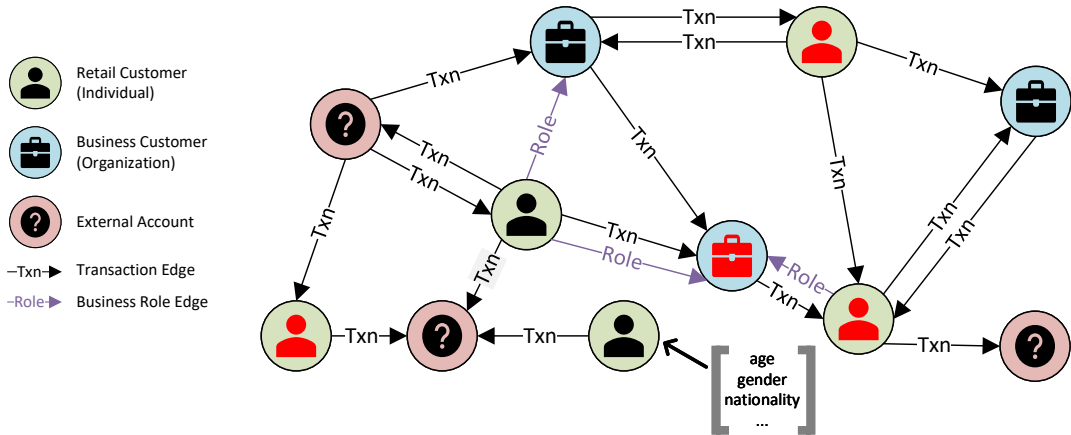


Vår graf

Homogen versus heterogen Graf



Homogen versus heterogen Graf



Grafens størrelse

Node-type	#Noder	#Variabler
Individer	1,718,159	11
Bedrifter	204,770	8
Eksterne	3,226,230	2
Totalt	5,149,159	—

Antall mistenkelige individer:
7,732 (0.45%)

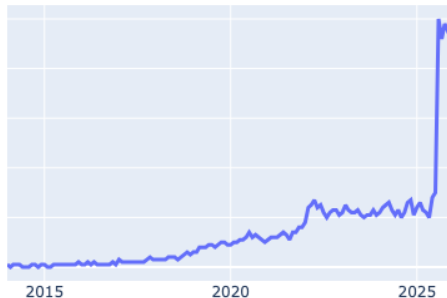
Kant-type	#Kanter	#Variabler
Transaksjoner	9,282,234	2
Roller	920,716	2
Totalt	10,202,950	—

Graph Neural Networks

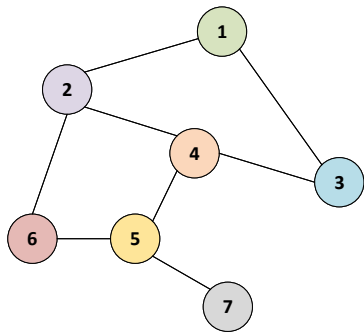
Graph Neural Network (GNN)

- Relativt nytt felt — vokst kraftig siden 2016
- Mange forskjellige algoritmer og rammeverk
- Sentral mekanisme: **Message Passing**

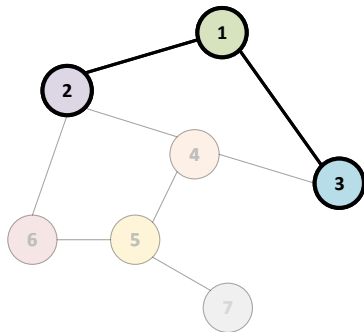
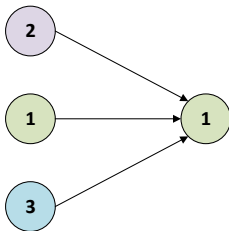
Google søk etter GNN



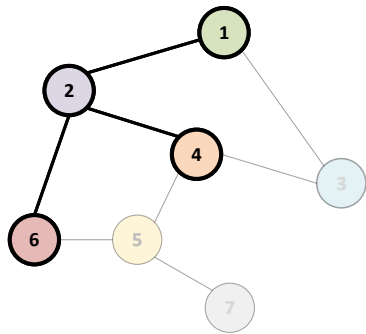
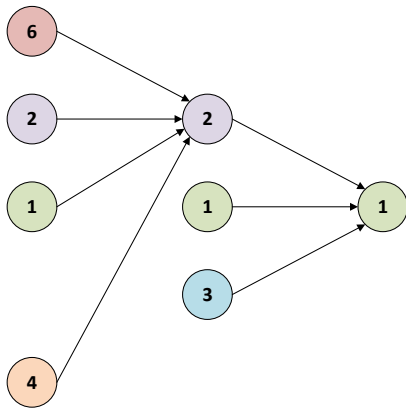
Message Passing



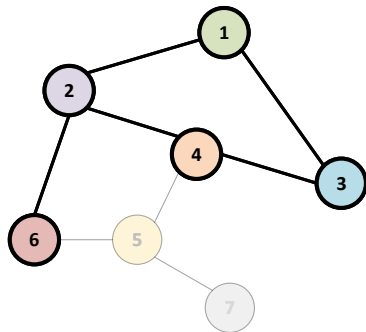
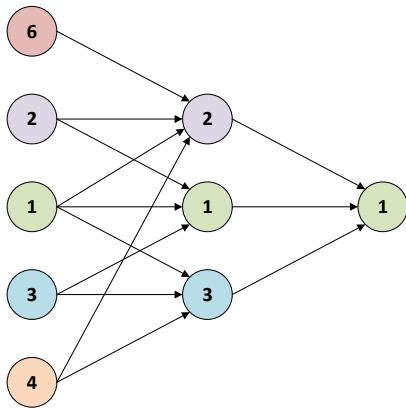
Message Passing



Message Passing

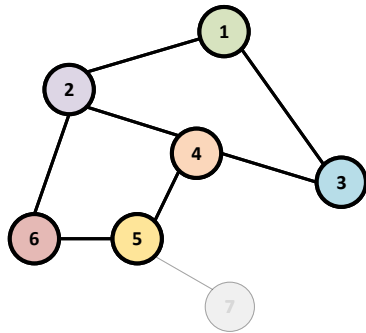
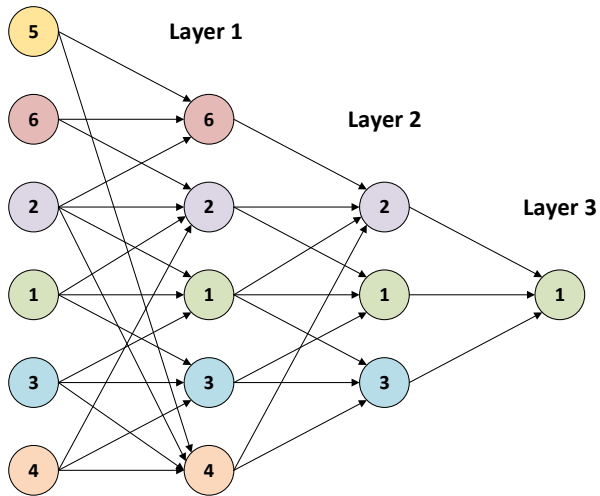


Message Passing



Message Passing

Layer 0



Graph Neural Network

De fleste eksisterende metodene fungerer kun på homogene grafer...

Vi trenger en metode som kan **utnytte rikheten i vår heterogene graf**:

- ? Håndterer flere nodetyper
- ? Håndterer flere kant-typer
- ? Utnytter variablene på både nodene og kantene

...metodeutvikling måtte til...

Løsning: Heterogeneous MPNN (H-MPNN)

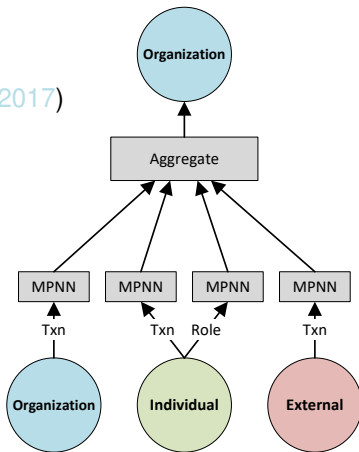
Vi **videreutvikler** den eksisterende metoden

Message Passing Neural Network (**MPNN**) ([Gilmer et al. 2017](#))

til **heterogene grafer**

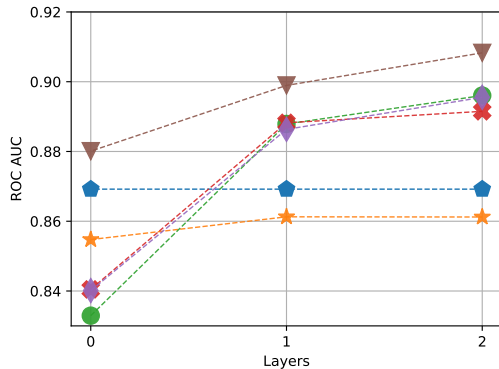
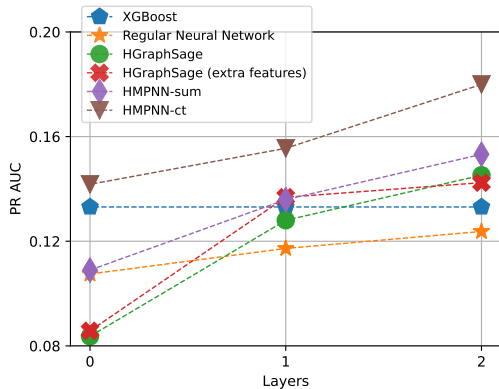
- ✓ Utnytter variablene på både noder og kanter
- ✓ Håndterer flere nodetyper
- ✓ Håndterer flere kant-typer

Software Library: PyTorch Geometric ([Fey et al. 2019](#))



Resultater

Resultater



Resultater



Resultater



Oppsummering

- Vi utforsket bruk av *Graph Neural Network* for å avdekke hvitvasking i en stor, heterogen graf basert på ekte data fra DNB
- Modellene ble trent ved bruk av veiledet læring for å predikere hvitvaskingsrisikoen til bankkunder
- GNN-rammeverket MPNN ([Gilmer et al. 2017](#)) ble utvidet til heterogene grafer
- Resultatene viser at GNN er en lovende metodikk for å avdekke hvitvasking



Fredrik Johannessen



**Bruk av grafbasert maskinlæring til
deteksjon av hvitvasking**

NR — Fagseminar om datadrevet
antihvitvasking og svindeldeteksjon