



UPPSALA
UNIVERSITET

Producera statistik från polisrapporter med AI

Måns Magnusson (Uppsala Universitet) och
Thomas Hvitfeldt (Brå)

6 Mars 2024



UPPSALA
UNIVERSITET

Text-som-data och statistik



- Text är centralt för att förmedla information



- Text är centralt för att förmedla information
- Två typer av "läsning":
 - Vanlig (ex. läsa **en** polisanmälan)



- Text är centralt för att förmedla information
- Två typer av "läsning":
 - Vanlig (ex. läsa **en** polisanmälan)
 - Statistisk "läsning" (ex. läsa **alla** polisanmälningar)



- Text är centralt för att förmedla information
- Två typer av "läsning":
 - Vanlig (ex. läsa **en** polisanmälan)
 - Statistisk "läsning" (ex. läsa **alla** polisanmälningar)
- Förstå viktiga **populationer** av dokument, exempelvis
 - Medicinska journaler
 - Domslut
 - Polisanmälningar



UPPSALA
UNIVERSITET

Hur hanteras textuell data idag?

- Text är komplext: "ostrukturerad data"



UPPSALA
UNIVERSITET

Hur hanteras textuell data idag?

- Text är komplex: "ostrukturerad data"
- Behöver ofta **klassificeras**, ex. anmälda hatbrott, typ av straff



Hur hanteras textuell data idag?

- Text är komplext: "ostrukturerad data"
- Behöver ofta **klassificeras**, ex. anmälda hatbrott, typ av straff
- Vanliga metoder för klassifikation idag:
 - Vid registrering



Hur hanteras textuell data idag?

- Text är komplext: "ostrukturerad data"
- Behöver ofta **klassificeras**, ex. anmälda hatbrott, typ av straff
- Vanliga metoder för klassifikation idag:
 - Vid registrering
 - Manuell kodning i efterhand



Hur hanteras textuell data idag?

- Text är komplext: "ostrukturerad data"
- Behöver ofta **klassificeras**, ex. anmälda hatbrott, typ av straff
- Vanliga metoder för klassifikation idag:
 - Vid registrering
 - Manuell kodning i efterhand
 - Textsökmetoder



Hur hanteras textuell data idag?

- Text är komplext: "ostrukturerad data"
- Behöver ofta **klassificeras**, ex. anmälda hatbrott, typ av straff
- Vanliga metoder för klassifikation idag:
 - Vid registrering
 - Manuell kodning i efterhand
 - Textsökmetoder
- **Kostsamt** med textuell data i **officiell** statistikproduktion



UPPSALA
UNIVERSITET

Sveriges Officiella Statistik



Sveriges officiella
statistik

- Statistik som används för
 - allmän information
 - utredningsverksamhet
 - forskning



UPPSALA
UNIVERSITET

Sveriges Officiella Statistik



Sveriges officiella
statistik

- Statistik som används för
 - allmän information
 - utredningsverksamhet
 - forskning
- Krav på officiell statistik:
 - objektiv
 - allmänt tillgänglig



UPPSALA
UNIVERSITET

Sveriges Officiella Statistik



Sveriges officiella
statistik

- Statistik som används för
 - allmän information
 - utredningsverksamhet
 - forskning
- Krav på officiell statistik:
 - objektiv
 - allmänt tillgänglig
 - korrekt
 - tillförlitlig



- Statistik som används för
 - allmän information
 - utredningsverksamhet
 - forskning
- Krav på officiell statistik:
 - objektiv
 - allmänt tillgänglig
 - korrekt
 - tillförlitlig
- Exempel:
 - Befolkningsstatistik (Statistiska Centralbyrån)
 - Dödsorsaksstatistik (Socialstyrelsen)
 - Kriminalstatistiken (Brottsförebyggande rådet)



- Statistik som används för
 - allmän information
 - utredningsverksamhet
 - forskning
- Krav på officiell statistik:
 - objektiv
 - allmänt tillgänglig
 - korrekt
 - tillförlitlig
- Exempel:
 - Befolkningsstatistik (Statistiska Centralbyrån)
 - Dödsorsaksstatistik (Socialstyrelsen)
 - Kriminalstatistiken (Brottsförebyggande rådet)
- Bygger ofta på textuella dokument



Transformerrevolutionen

- Transformer-modellen

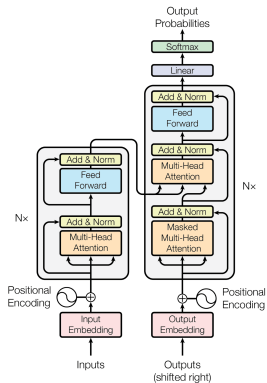


Figure:

Transformermodellen
(Vaswani et al., 2017)



Transformerrevolutionen

- Transformer-modellen
- Möjliggjort **förtränade** språkmodeller (ex. chatGPT)

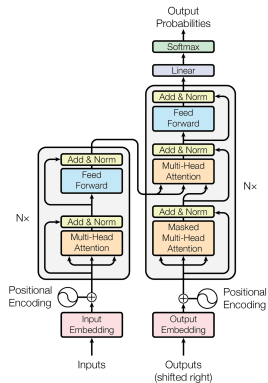


Figure:

Transformermodellen
(Vaswani et al., 2017)



Transformerrevolutionen

- Transformer-modellen
- Möjliggjort **förtränade** språkmodeller (ex. chatGPT)
- Mycket bra för automatisk klassifikation av text

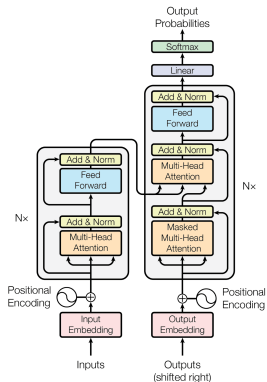


Figure:
Transformermodellen
(Vaswani et al., 2017)



Transformerrevolutionen

- Transformer-modellen
- Möjliggjort **förtränade** språkmodeller (ex. chatGPT)
- Mycket bra för automatisk klassifikation av text
- Kan anpassas till olika texter

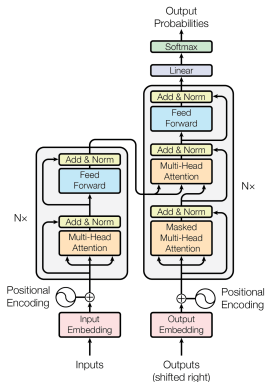


Figure:

Transformermodellen
(Vaswani et al., 2017)



UPPSALA
UNIVERSITET

Kan vi använda oss av transformermodeller för att producera statistik?



UPPSALA
UNIVERSITET

Vårt mål: Steg 1

- Vårt mål: **Träna en hatbrottsmodell**

"Ej hatbrott" = `transformer("Min cykel har blivit stulen.")`



UPPSALA
UNIVERSITET

Vårt mål: Steg 1

- Vårt mål: **Träna en hatbrottsmodell**
"Ej hatbrott" = `transformer("Min cykel har blivit stulen.")`
- Tränar på kända data från Brå



UPPSALA
UNIVERSITET

Vårt mål: Steg 1

- Vårt mål: **Träna en hatbrottsmodell**

"Ej hatbrott" = $\text{transformer}(\text{"Min cykel har blivit stulen."})$

- Tränar på kända data från Brå
- Klassificerar okända data (anmälningar)



Vårt mål: Steg 1

- Vårt mål: **Träna en hatbrottsmodell**
"Ej hatbrott" = $\text{transformer}(\text{"Min cykel har blivit stulen."})$
- Tränar på kända data från Brå
- Klassificerar okända data (anmälningar)
- Räkna anmälningar som klassificeras som hatbrott



UPPSALA
UNIVERSITET

Data och metoder

- Data
 - Samtliga polisanmälningar sedan 2007:
cirka **17.7 miljoner anmälningar**



- Data
 - Samtliga polisanmälningar sedan 2007:
cirka **17.7 miljoner anmälningar**
 - Alla manuellt bedömda hatbrott fram till 2020:
52 062 konstaterade hatbrott
60 749 konstaterade icke-hatbrott



- Data
 - Samtliga polisanmälningar sedan 2007:
cirka **17.7 miljoner anmälningar**
 - Alla manuellt bedömda hatbrott fram till 2020:
52 062 konstaterade hatbrott
60 749 konstaterade icke-hatbrott
- Metod
 - KB-BERT (Malmsten et al, 2020)



- Data
 - Samtliga polisanmälningar sedan 2007:
cirka **17.7 miljoner anmälningar**
 - Alla manuellt bedömda hatbrott fram till 2020:
52 062 konstaterade hatbrott
60 749 konstaterade icke-hatbrott
- Metod
 - KB-BERT (Malmsten et al, 2020)
 - Polis-BERT



Prel. resultat och utmaningar

Sann/klassifikation	Hatbrott	Ej hatbrott
Hatbrott	3307	91
Ej hatbrott	135	2497
Ej bedömda	13190	1475814

Table: Modell 1 för år 2020 (96,3% korrekt)

- Utmaningar
 - Hur kan detta användas? **Modellen överskattar hatbrott.**



Prel. resultat och utmaningar

Sann/klassifikation	Hatbrott	Ej hatbrott
Hatbrott	3307	91
Ej hatbrott	135	2497
Ej bedömda	13190	1475814

Table: Modell 1 för år 2020 (96,3% korrekt)

- Utmaningar
 - Hur kan detta användas? **Modellen överskattar hatbrott.**
 - Osäkerhet och bias i klassifikationerna



Prel. resultat och utmaningar

Sann/klassifikation	Hatbrott	Ej hatbrott
Hatbrott	3307	91
Ej hatbrott	135	2497
Ej bedömda	13190	1475814

Table: Modell 1 för år 2020 (96,3% korrekt)

- Utmaningar
 - Hur kan detta användas? **Modellen överskattar hatbrott.**
 - Osäkerhet och bias i klassifikationerna
 - Kombinera manuell läsning, statistiska urval och språkmodeller



UPPSALA
UNIVERSITET

Tillämpningar vid Brå

brå