

Artificial Intelligence Support in Medical Emergency Calls AISMEC

Guttorm Brattebø
Seksjonsoverlege / leder KoKom
Kirurgisk Serviceklinikk, Haukeland universitetssykehus
Professor, K1, Universitetet i Bergen

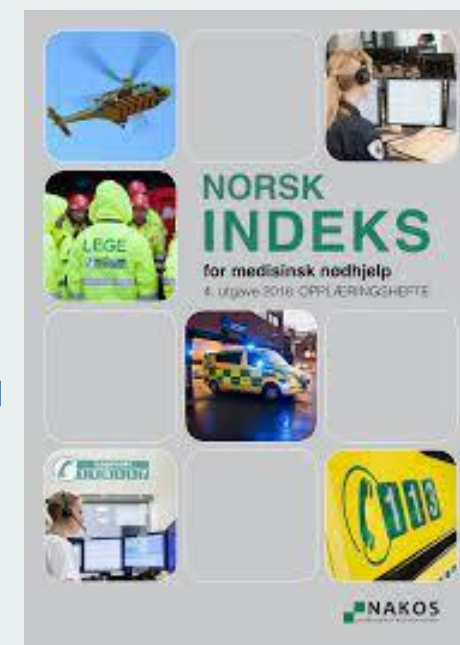
På vegne av AISMEC-prosjektet



Problemet

- Hvordan skal AMK identifisere de som har tegn på hjerneslag?
- AMK kan sende ut ambulanse eller luftambulanse
- Uendret forekomst av slag, men økende antall innlagte
- Noen «faller mellom stolene» (ca 10%)
- Hvordan klare å få oversikt via telefon?





Symptomdebut



Mulig løsning

- De fleste av oss har historiske data i spesialisthelsetjenesten som er «kategoriske»
- Er det mulig å beregne en sannsynlighet for slag / ikke-slag basert på dette?
- Kan AI kombinere tidligere sykehistorie fra AMIS, EPJ, prosedyrer, resepter, blodprøvesvar og samtalen med AMK?



Datakilder & samarbeidspartnere

- Data fra 2 år i Helse Bergen
- Pasienter identifisert i NHR
- AMIS
- DIPS
- Blodprøveresultater
- Legemiddelregisteret
- KoKom, KSK, HUS
- NHR
- Nevroklinikken, HUS
- Helsetjenesteutvikling, HUS
- OUS
- HVL / NHH
- LHL / Hjerneslag
- HVIKT



AI-modellen

- Automatisk tale-tekst gir transkripsjonsmuligheter
- Nødsamtaler er vanskelige: aksenter, bakgrunnsstøy og spesialisert vokabular
- Vi startet med Whisper-modellen fra Nasjonalbiblioteket (NB), forhåndstrent med et omfattende norsk språk-korpus, inkludert ulike regionale dialekter
- Vi finjustert modellen basert 86 manuelt transkriberte lydfile med nøyaktig ekspert-annotering for å sikre høy nøyaktighet
- Datasettet ble delt i tre: 80 % for trening (n=69), 10 % for validering og 10 % for testing
- Godt grunnlag for å tilpasse modellen til de spesifikke lydbildene i nødsamtalene



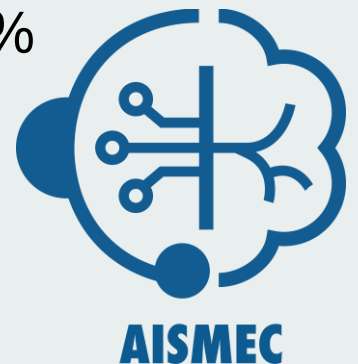
Mer om AI-modellen

- Modellens ytelse ble vurdert med Word Error Rate (WER)
- Den opprinnelige NB-Whisper-modellen gav ordfeil-rate på 0,83 på test-datasettet, mens den finjusterte modellen hadde relativ forbedring på 5% (WER 0,78), noe som indikerer en betydelig forbedring nøyaktighet
- Transkribert tekst ga input til en modell for å anslå sannsynlighet for hjerneslag
- Vi valgte Modern-BERT-arkitekturen for å håndtere lange sekvenser, fordi nødsamtaler kan variere i lengde, og ofte overstige grensen for kontekstvinduet til standard BERT-baserte modeller
- Modern-BERTs utvidede kontekstlengde (8192 tokens) gjør det mulig å inkludere hele samtaletranskripsjoner uten å avkorte, noe som bevarer potensiell verdifull informasjon
- Den overordnede pipeline-strukturen er som følger: Nødsamtale (lyd) > Whisper > transkripsjon > Modern-BERT > klassifisering
- Resultatet av Modern-BERT-modellen er en sannsynlighets-score for slag
- Skalaen er fra 0 til 1, hvor høyere verdier indikerer større sannsynlighet for slag



Resultater

- Sammenstilling av data - krevende fra mange kilder
- Lyd til tekst-modul fungerer bra
- AI-modellen presterer også bra
 - Sensitivitet 81% (95%CI: 78–84)
 - Spesifisitet 80% (95%CI: 79–84)
 - Positiv prediktiv verdi PPV 54% (95%CI: 51–57)
 - Negativ prediktiv verdi NPV 93% (95%CI: 92–94)
- En dansk AI-basert modell for å identifisere hjertestans basert på over 340.000 samtaler, rapporterte en sensitivitet på 63% og en PPV på 25%



Neste skritt

- Teste ut algoritmen på et uselektert materiale fra AMK (ca 50.000 nødsamtaler)
- Foreløpig avslag fra REK...
- Om dette funker like godt; tilgjengeliggjøre for å kjøre «live» i AMK
- Mer info: www.helse-bergen.no/aismec
- Takk til NFR og Lærdalsfondet for finansiering

