

Fra lyd til tekst

Gruppe 12

Guro Kalland

Lene Odde

Martin Nenseth

Sindre Grønstøl Haugeland

Thomas Berntsen

Robert Berntsen

Innlevering denne gangen skal være en kort artikkel (husk referanser!) + en kort presentasjon.

De aller første talegjenkjenningsprogrammene på 50 tallet klarte kun å forstå tall. I 1962 laget IBM [Shoebox](#) som forstod 16 engelske ord. De siste 15 årene har utviklingen av talegjenkjenningsapplikasjoner, både for PC, nettbrett og smarttelefoner, eksplodert. Skype jobber nå med en ny versjon som vil [oversette språk i tilnærmet sann tid](#). Store kjemper i markedet som Microsoft, Apple og Google er enige om at det finnes stort potensial i talegjenkjenningsprogrammer, og at det *kan endre måten vi bruker enhetene våre på i fremtiden*.

Siri

Apple sin Siri er utvilsomt et godt eksempel på taleforståelse. Det er viktig at datasystemer forstår spørsmålet for å kunne finne riktig svar.

Slike spørresystemer har et tale-til-tekst-system som gir grunnlaget for spørringen til flere kunnskapsdatabaser. Apple spør videre til Google, Yelp, Wolfram Alpha, Yahoo og Wikipedia. Etter at dette er blitt gjort sammenstilles svaret til det som gir høyest statistisk relevans i forhold til hvordan spørsmålet tolkes. Selv om Siri er en av de beste på taleforståelse, forstår den ikke alt.

Men det er flere eksempler på taleforståelse som blir brukt, for eksempel Google Now, Microsoft Cortana osv. Disse kan fungere som "personlig assistanse" som lytter til det du sier, finner ut hva du mener for å så å forsøke å gjøre hva du spør etter. dette kan være å sette på en alarm til et gitt tidspunkt, finne nærliggende restauranter, spørre om telefonnummer osv.

Google's voice recognition

Med Androids oppgradering til Jelly Bean ble også Googles talegjenkjenningssystem sterkt oppgradert, brukernes feilmeldinger ble redusert med 25%.

Google implementerte et "nervesystem" som deler opp det brukeren sier i åtte biter, og

sender bitene til Googles databaser verden over for å prosessere det. Før brukeren får noen tilbakemelding analyserer Google mange forskjellige mønstre, skiller konsonanter og vokaler, og avgir et intelligent svar til brukeren.

Det er fire måter et talegjennkjenningssystem bruker for å kjenne igjen tale på;

- 1) Ved bruk av enkel mønstergjennkjenning. Der maskinen gjenkjenner ord i sin helhet ut ifra et tidligere observert mønster.
- 2) Mønster og funksjon analyse - det er der hvor er delt i biter og anerkjent fra viktige funksjoner. De fleste av oss har store vokabular, vi har hundrevis av ord lagret
- 3) Modellering av språk og statiske analyse - hvor kunnskap om grammatikk og sannsynlighet for hyppighet av visse ord, brukes til at programvaren kan gjette til seg de ord du faktisk mener raskere og mer nøyaktig.
- 4) Nevral nettverk - kan gjenkjenne mønstre, som for eksempel ord lyder.

Referanser:

<http://www.explainthatstuff.com/voicerecognition.html>

http://www.phonearena.com/news/The-secret-of-Googles-amazing-voice-recognition-revealed-it-works-like-a-brain_id39938

<http://www.tu.no/artikler/slik-kan-datamaskinene-snakke-og-forsta-tale/234230>