ICA – Uke 5

# Oppgave 1.2.1 – Huffman code

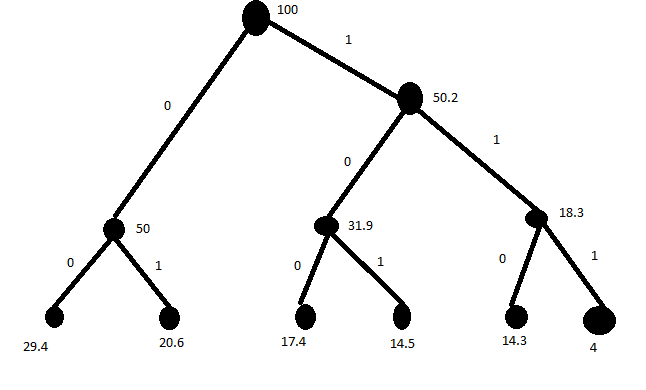
## A)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| UiA fakultet | Helse og idrett | Humanoira og ped | Kunstfag | Tech og realfag | Lærerutdanning | Økonomi og samfunn |
| Antall studenter | 1829 | 1525 | 420 | 2166 | 1506 | 3093 |
| % sannsynlighet | 17.4% | 14.5% | 4.0% | 20.6% | 14.3% | 29.3% |
| Sannsynlighet | 0.17 | 0.15 | 0.04 | 0.21 | 0.14 | 0.29 |

## B) Minst informasjon for hvilket fakultet

Man får minst informasjon om Økonomi og Samfunnsfag, fordi det er 29.3% sannsynlighet for at en student tilhører dette fakultetet.

## C) Huffman Tre



Økonomi og samfunn – 00 – 2 bits

Tech & realfag – 01 – 2 bits

Helse og idrett – 100 – 3 bits

Humanoira & ped – 101 – 3 bits

Lærerutdanning – 110 – 3 bits

Kunstfag – 111 – 3 bits

## D) Gjennomsnittslengde for melding som inneholder fakultetskoder for 100 tilfeldig valgte studenter er 250.

Fordi:

Vi ganger antall bits for hvert fakultet med sannsynligheten for at studenten tilhører det fakultetet.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| UiA fakultet | Helse og idrett | Humanoira og ped | Kunstfag | Tech og realfag | Lærerutdanning | Økonomi og samfunn |
| Antall studenter per fakultet | 1829 | 1525 | 420 | 2166 | 1506 | 3093 |
| Sannsynlighet | 0.17 | 0.15 | 0.04 | 0.21 | 0.14 | 0.29 |
| Bits | 3 | 3 | 3 | 2 | 3 | 2 |

Det er hundre studenter totalt, derfor må vi gange antall bits med sannsynlighet for at en student hører til ett spesifikt fakultet for å finne ut totalt antall bits.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sannsynlighet | 0.17 | 0.15 | 0.04 | 0.21 | 0.14 | 0.29 |
| Bits | 3 | 3 | 3 | 2 | 3 | 2 |
| Bits \* sannsynlighet | 0.51 | 0.45 | 0.12 | 0.42 | 0.42 | 0.58 |
| Antall bits per fakultet når 100 tilfeldige studenter er valgt | 51 | 45 | 12 | 42 | 42 | 58 |

Gjennomsnittslengden for en melding blir da 250 bits.

Entropi:

0.29\*log2(1/0.29) = 0.5179038064

0.21\*log2(1/0.21) = 0.4728231410

0.17\*log2(1/0.17) =0.43458686924

0.15\*log2(1/0.15) = 0.4105448391

0.14\*log2(1/0.14) = 0.3971101774

0.04\*log2(1/0.04) = 0.1857542475

Svar: 243 bits

Ved bruk av entropi i utregning får vi noen færre bits enn ved bruk av Huffman metoden. Dette er fordi Huffmann metoden ikke er helt optimal når det kommer til komprimering. Men det er en enkel måte å regne ut ca bits på.

## E) Se i python fil oppgave E