

AI i prosjekter og prosjektledelse

KunnskapsArena

Bruk av maskinlæring for å finne kjennetegn på prosjektsuksess.

Eksempel fra bruk av Nordic 10-10 data.



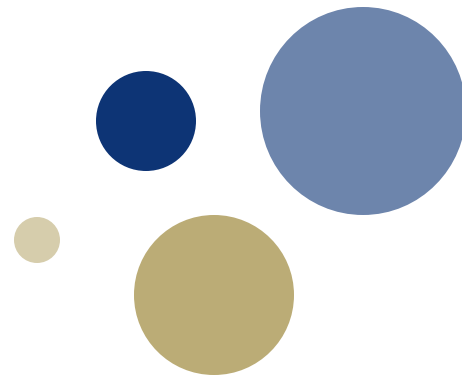
Bruk av maskinlæring

- Samler og analyserer prosjektbaserte dataset
 - Evaleringer
 - Planer, store planer/prosjekt, gjerne flere baselines
 - Oppfølging av prosjekter, gjerne flere prosjekter
 - Pågår forskningsprosjekter med bruk av eksempelvis data fra anleggsmaskiner
- Evalueringer, små dataset men forholdsvis konsistente



NTNU


Kunnskap for en bedre verden



Bruk av maskinlæring for å finne kjennetegn på prosjektsuksess. Eksempel fra bruk av Nordic 10-10 data

Hva kan hentes ut ved bruk på mindre datasett

Av: Magnus Aarvold og Wilhelm Hartvig



Innhold for presentasjonen

- Kort om oss og oppgaven
- 10-10 datasettet: hva, hvorfor, hvordan
- Kort om metode
- Hvordan definere suksess bestemmer modellen
- Resultater fra analysen

Wilhelm Jan Hartvig

- Befalsskole
- NTNU
- Prosjektledelse og ny, avansert teknologi
- Skal jobbe som konsulent i krysningen mellom prosjektledelse og IT



Magnus Olai Aarvold

- NTNU
- Ny teknologi og innovasjon
- Skal jobbe som utvikler



Oppgaven

- Bruke maskinlæring på prosjektdata for å se hva som kan hentes ut

CII Nordic 10-10 datasettet – bakgrunn

- Laget for å være et "benchmark" i industrien
- Verktøy for diskusjon og forbedringer
- Bevist at deltagende bedrifter yter bedre enn andre



10-10 data – oppsett

- Exportert fra SPSS
- G – generell
- I – input
- O – output
- Q – questionnaire

O_01	O_02	O_03	O_04	O_05	O_06
-0,15	-0,06	143,45	0,04	#NULL!	#NULL!
-0,05	#NULL!	86,19	0,04	#NULL!	#NULL!
0,02	0,00	33,33	0,08	#NULL!	#NULL!
-0,45	0,02	0,59	0,05	0,10	#NULL!
-0,12	0,00	7,41	0,39	#NULL!	#NULL!
0,07	0,04	1,08	#NULL!	#NULL!	#NULL!
0,33	0,00	5,04	0,03	0,07	#NULL!
0,10	0,06	22,10	0,03	0,02	#NULL!
-0,11	-0,36	8,32	0,02	0,01	#NULL!
-0,12	0,00	3,43	0,07	0,11	#NULL!
0,00	-0,04	1,21	#NULL!	#NULL!	#NULL!
-0,11	-0,04	17,35	0,30	#NULL!	#NULL!
0,10	0,30	3,58	#NULL!	#NULL!	#NULL!
0,00	0,02	3,60	0,06	0,06	#NULL!

Q001a	Q001b	Q001c	Q001d	Q001e	Q001f	Q001g
0,13	0,63	0,63	0,38	0,00	0,00	0,13
0,08	0,31	0,15	0,08	0,15	0,69	0,08
0,13	0,25	0,13	0,38	0,38	0,38	0,00
0,00	0,00	0,00	0,00	0,13	0,63	0,13
0,00	0,06	0,18	0,18	0,00	0,47	0,06
0,00	0,11	0,11	0,00	0,11	0,44	0,00
0,00	0,60	0,10	0,10	0,20	0,30	0,10
0,07	0,40	0,13	0,33	0,27	0,40	0,07
0,00	0,69	0,19	0,13	0,56	0,31	0,00
0,00	0,57	0,07	0,21	0,21	0,64	0,24
0,00	0,00	0,00	0,08	0,08	0,50	0,00
0,00	0,00	0,00	0,00	0,20	0,60	0,00

I1_Pr	I2_Pr	I3_Pr	I4_Pr	I5_Pr	I6_Pr	I7_Pr
71	72	67	67	76	66	65
71	73	73	65	71	65	65
76	69	70	68	78	64	68
72	76	76	70	69	76	70
67	64	64	60	65	55	59
84	80	79	82	94	74	86
71	63	68	61	62	57	62
71	67	74	73	71	71	69
70	70	67	69	71	70	70
69	62	62	65	63	61	65
85	81	79	80	77	82	83
92	87	89	85	98	76	89
64	65	63	65	68	66	63
68	71	69	69	68	77	69
87	84	86	82	92	69	93

G1_Prokje	G2_Prokje	G3_Prokje	G4_Kontra	G5_Antall	G6_Flere_
3	4	1	2	8	0
3	4	4	2	13	0
3	4	1	4	8	1
3	2	4	1	8	0
3	4	4	4	17	1
3	1	4	2	9	0
3	2	1	1	10	0
3	2	1	1	15	0
3	2	1	4	16	0
3	3	1	4	14	0

10-10 datasettet – innhold

- Spørsmål rettet til ulike faser og sektorer

Variabel SPSS	Construction					Infrastructure					Industry				
	Phase 0	Phase 1	Phase 2	Phase 3	Phase 4	Phase 0	Phase 1	Phase 2	Phase 3	Phase 4	Phase 0	Phase 1	Phase 2	Phase 3	Phase 4
	spm 9	spm 9	spm 8	spm 9	spm 8	spm 8	spm 8	spm 7	spm 8	spm 6	spm 8	spm 8	spm 7	spm 8	spm 6
Q007a	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Q007b	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Q007c	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Q007d	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Q007e	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Q007f	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	spm 10					spm 9					spm 9				
Q008a	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
Q008b	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
Q008c	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
	spm 11	spm 10	spm 9			spm 10	spm 9	spm 8			spm 10	spm 9	spm 8		
Q009a	1	1	1	0	0	1	1	1	0	0	1	1	1	0	0
Q009b	1	1	1	0	0	1	1	1	0	0	1	1	1	0	0

Kort introduksjon til maskinlæring

- Økning i popularitet fra 2000
- Predikere suksess er utbredt
- State-of-the-Art
 - Automated Project Management
 - Smart erfaringsoverføring
 - IoT med tilhørende programvare som hjelp til prosjektleder

Definere suksess

Løsning A

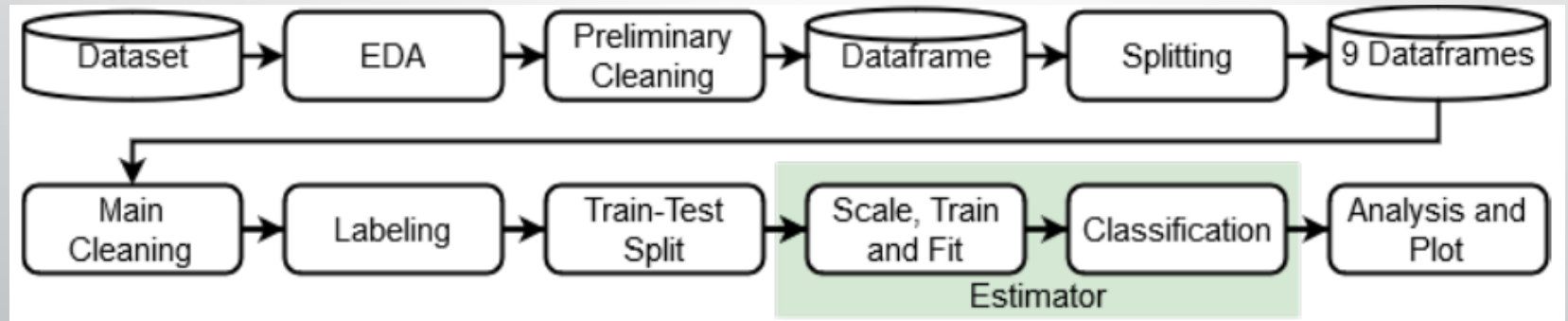
- Legge sammen selekterte verdier som gir én samlet score for prosjektet
- Hvis: score > 0 → Success
- Hvis: score < 0 → Failure

Solution	Success	Failure	Total
A	51	30	81
A_fillNA	75	49	124
B	68	56	124

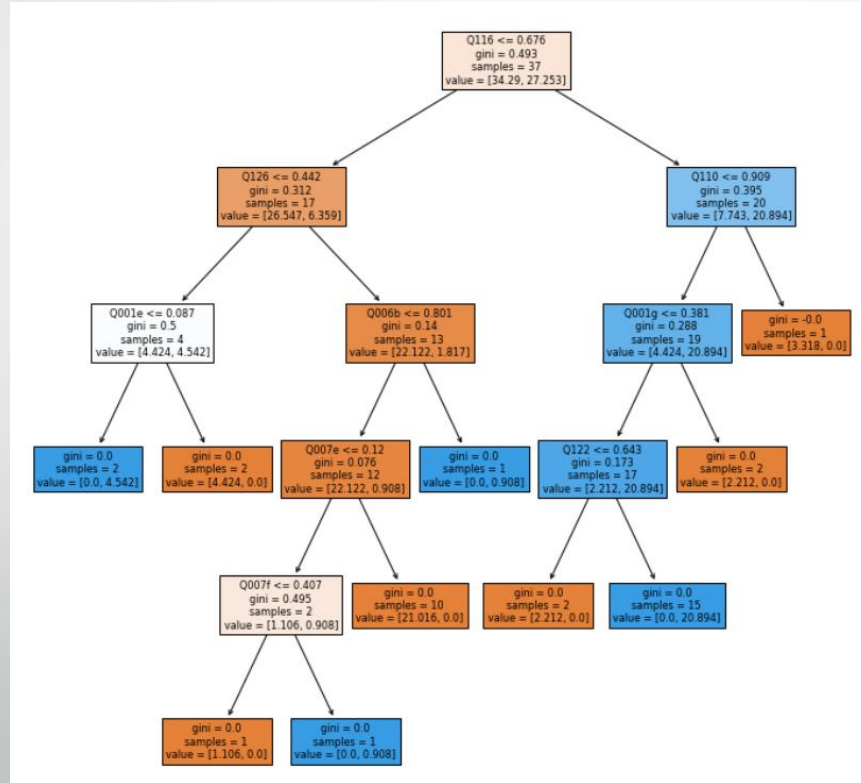
Løsning B

- 2-ut-av-3
- Kost, tid og kundetilfredshet
- Hvis: 2 eller flere verdier er ok er det suksess
- Ellers: prosjektet er "failure"

Metode – Oversikt



Metode - Random Forrest Classifier



Vår modell

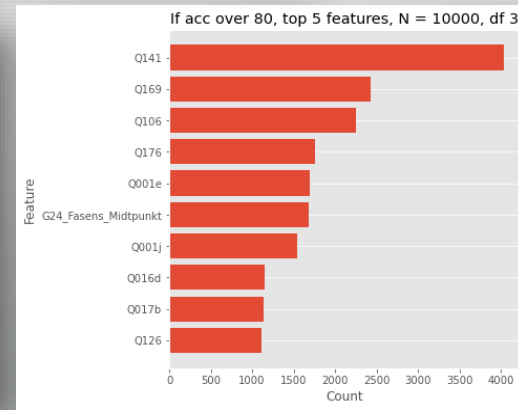
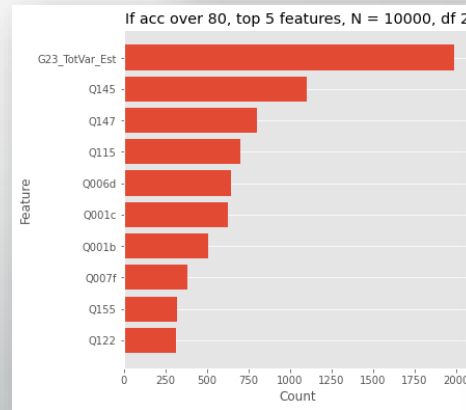
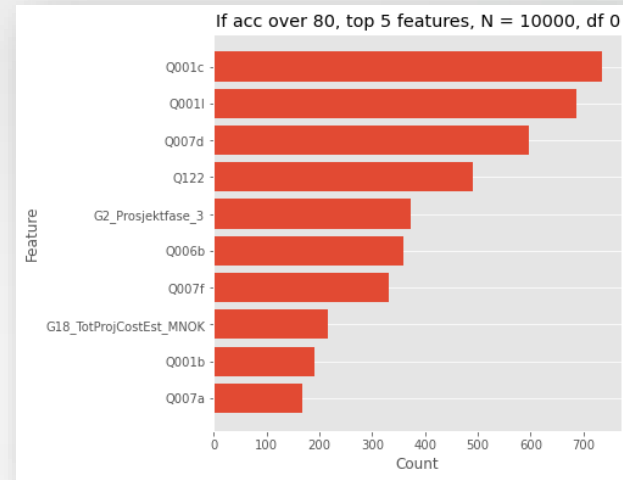
- Dele inn datasett basert på sektor og fase
 - 9 sett hvor 4-5 av disse oppnådde god nøyaktighet

- Rense og fyller felter
- Monte Carlo

DF#	Split	Shape (row, col)	Shape after NaN drop (row, col)	# success	% success	Mean Accuracy	Mean F1-score	Mean CM
0	Whole construction sector	(47, 296)	(47, 51)	28	60	0.54	0.62	[[1.5 2.5] [2.1 3.9]]
1	Whole infrastructure sector	(77, 296)	(77, 38)	41	53	0.7	0.72	[[4.8 2.2] [2.66 6.34]]
2	Construction phase 1	(22, 296)	(22, 86)	12	55	0.55	0.62	[[0.75 1.25] [1.01 1.99]]
3	Construction phase 3	(16, 296)	(16, 105)	12	75	0.69	0.8	[[0.01 0.99] [0.26 2.74]]
4	Infrastructure phase 1	(32, 296)	(32, 79)	14	44	0.73	0.64	[[3.32 0.68] [1.17 1.83]]
5	Infrastructure phase 3	(28, 296)	(28, 99)	20	71	0.73	0.81	[[0.83 1.17] [0.45 3.55]]
6	Both sectors only phase 0	(17, 296)	(17, 80)	7	41	0.48	0.05	[[1.82 0.18] [1.92 0.08]]
7	Both sectors only phase 1	(54, 296)	(54, 70)	26	48	0.67	0.64	[[3.93 2.07] [1.6 3.4]]
8	Both sectors only phase 3	(44, 296)	(44, 92)	33	75	0.67	0.78	[[0.45 1.55] [1.44 5.56]]

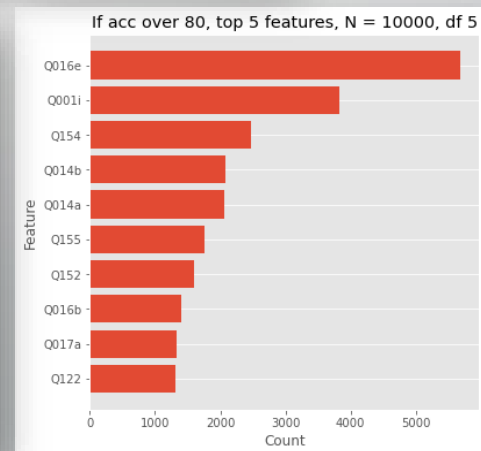
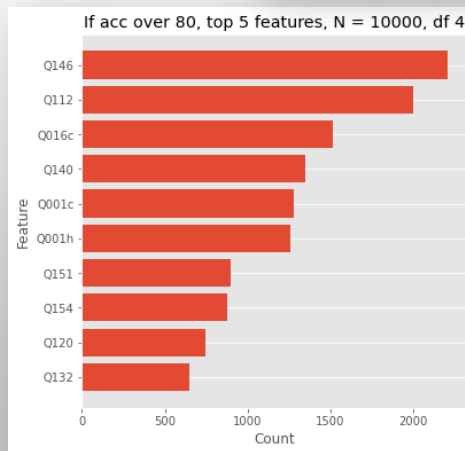
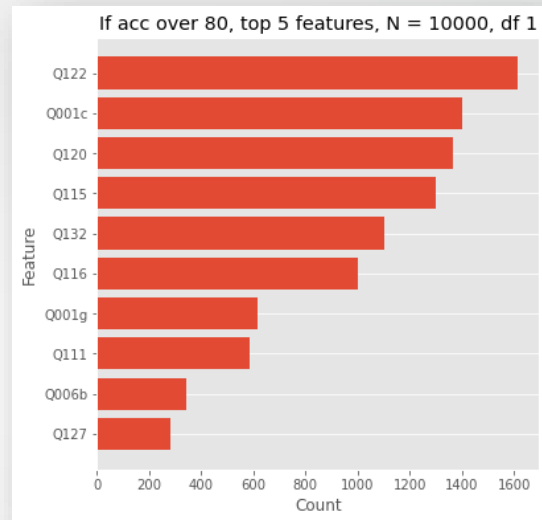
Bygg & Anlegg

- Q001c: fremdriftsplan → komplekst
- Q001l: fortettet arbeidsområde → komplekst
- Q007d: avgørelser ble umiddelbart kommunisert
- Q122: prosesser støtter suksess
- Q141: endringsledelse
- Q145: kvalitetsstyring



Infrastruktur

- Q001c: fremdriftsplan → komplekst
- Q016c: endringer i listen over hovedkomponenter
- Q016e: stort antall avviksmeldinger
- Q112: strategi og plantidlig kommunisert til PL-team
- Q115: usikkerhetsvurdering
- Q120: involvering av byggherre
- Q122: prosesser støtter suksess
- Q132: profesjonell opplæring
- Q146: forslag til forbedring



Resultater

Table 5.1: Top features from the best performing dataframes with their conceptual meaning.

Feature	Description	Concept	DF
- Q001c	The complexity was very high due to the progression plan	Complexity	1, 4, 7
- Q016c	The project had a large quantity of changes in the list of main components	Changes	4
- Q016e	The project had a large quantity of deviation reports	Deviations	5
+ Q112	The tender plan were developed and communicated to the project team during the engineering phase	Planning	4, 7
+ Q115	All necessary and relevant members of the project team were involved in the process for uncertainty analysis	Uncertainty	1, 7
+ Q120	The involvement from project owner were appropriate	Leadership involvement	1, 4
+ Q122	The projects processes and systems supports project success	Project owner's process	1
+ Q132	Members of the project team participated an adequate professional training directly associated with their engineering work	Training	1
+ Q146	A rich supply of suggestions for improved constructability were evaluated and integrated during the engineering phase	Planning	4, 7
+ Q147	Cost to fix potential faults were considered during the engineering phase	Cost of quality	7

DF1 hele infrastruktur – DF4 infrastruktur fase 1 – Df7 begge sektorer fase 1

Faktorer som ikke er blant toppen

- Ingen faktorer som nevner BIM (av tre mulige)
- Q013: Ble prosjektets hovedmål endret?
- Q105: Interessenter ble håndtert på en god måte
- Q111: Høy andel tillit, respekt, åpenhet mellom aktørene?
- Q114: Nøkkemedlemmer forstod eiers mål og hensikt med prosjektet
- Q126: Ledergruppen kommuniserte strategiske mål, prosjektets mål på en effektiv måte

Utfordringer

- Små datasett
- Ufullstendige datasett

Hva er viktigst for dere?

- Struktur, struktur, struktur
- Fullstendig datasett
- Mye data

Spørsmål

- Faktorer som ikke er på toppen, hvorfor?
- Refleksjoner om at maskinlæring finner faktorer som predikerer, men ikke om det er for suksess eller fiasko
- Har dere brukt maskinlæring på prosjektbaserte datasett?
- Hvilke data kan være aktuelle for noe lignende?



Takk for oss!