

Prediksjonsmarkeder: Oppdaterte erfaringer

Stein Grimstad (steingr@simula.no)

Simula Research Laboratory & Wasteless

[**simula** . research laboratory]

wasteless

Forskningsmateriale

- ✓ To felt-studier
- ✓ En survey
- ✓ Litteratur-studie



PREDIKSJONSMARKEDER

- ✓ Prediksjonsmarkeder er verktøy som gjør det praktisk mulig å få tilgang til, samt aggregere, all estimerings-relevant informasjon i organisasjonen
- ✓ Prediksjonsmarked implementeres som elektroniske markeder, der alle ansatte kan delta, og hvor premiering er koblet til fremtidige hendelser

The screenshot shows the Intrade website interface. At the top, the logo "intrade" is displayed with the tagline "The Prediction Market". A quote from Kolemian Strumpf, Professor of Economics at the University of Kansas, is visible: "When people ask me who will win the next election, I say 'Let me look at the price on Intrade. It is the best forecast I know.'" The navigation menu includes Home, Signup, Markets, Trade, Practice, and Forum. A "READY TO START?" section features a prominent "Open An Account" button. Below this, there is a "Welcome Video" section with a play button icon and a "More Videos" link. Further down, there are links for "Introduction & Example Trade", "Register For a Webinar", "Intrade News", and "PDA Interface". The main content area displays a "MARKET SEARCH:" field with a "Search" button. Below the search field, the prediction market title "Will project AcmeWeb release to prod before 24.12.2008?" is shown. Underneath, the "Selected Prediction Markets" section is visible. The market details show a "Last Price: 61.2" and a "Closing Price: 61.4". There are "Buy" and "Sell" buttons, along with "You can buy this at 61.2" and "You can sell this at 61.1" text. A line chart shows the price movement from 2007 to 2008. At the bottom, there are "Explain" and "Trade Now" buttons.

Anvendelse av prediksjonsmarked

Markedet (spørsmålet)

- Et tidsavgrenset spørsmål relatert til en fremtidig hendelse eller parameterverdi
- Eksempel: "Når blir vårt nye billettsystem satt i produksjon?"

Kontraktene (svaralternativene)

- Én kontrakt representerer ett bestemt utfall
- Eksempel på kontrakter:
A:[innen uke 46], B:[uke 47-50], C:[etter uke 50]

Markedsplassen (applikasjonen)

- En webapplikasjon hvor deltakerne kjøper og selger kontrakter
- Markedsplassen genererer kjøps- og salgspriser basert på handlene
- Deltakerne kjøper det de synes er billig, og selger når de får en god pris

Hvordan prises kontraktene?

Prisen reflekterer sannsynligheten for at hendelsen tilknyttet kontrakten inntreffer.

Eksempel:

Når blir vårt nye billettsystem satt i produksjon?

A:[innen uke 46], B:[uke 47-50], C:[etter uke 50]

Pris på de ulike kontraktene (kan for eksempel være):

A: 16 kroner = 16% sannsynlighet

B: 27 kroner = 27% sannsynlighet

C: 57 kroner = 57% sannsynlighet

HVA SKJER NÅR OLE HANDLER?

$$K(A,B,C) = b * \ln(e^{A/b} + e^{B/b} + e^{C/b})$$

$$P(C) = b * e^{C/b} / (e^{A/b} + e^{B/b} + e^{C/b})$$

- Prisene på alle kontraktene oppdateres etter hver handel (av en automatisert "market maker")
- Hvis Ole f eks kjøper 80 aksjer i kontrakt C, vil C stige i kurs mens A og B vil synke (slik at summen fortsatt er 100)
- Ole betaler $K(A,B,C+80) - K(A,B,C)$ for aksjene sine

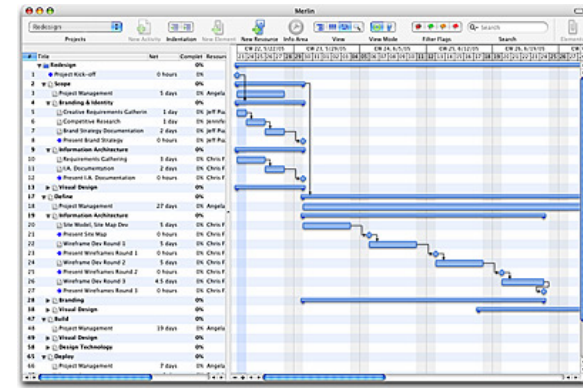
Prediksjonsmarkeder har blitt brukt til estimering, med suksess, i flere IT-prosjekter (Google, Siemens, IBM, ...)



Antall feil

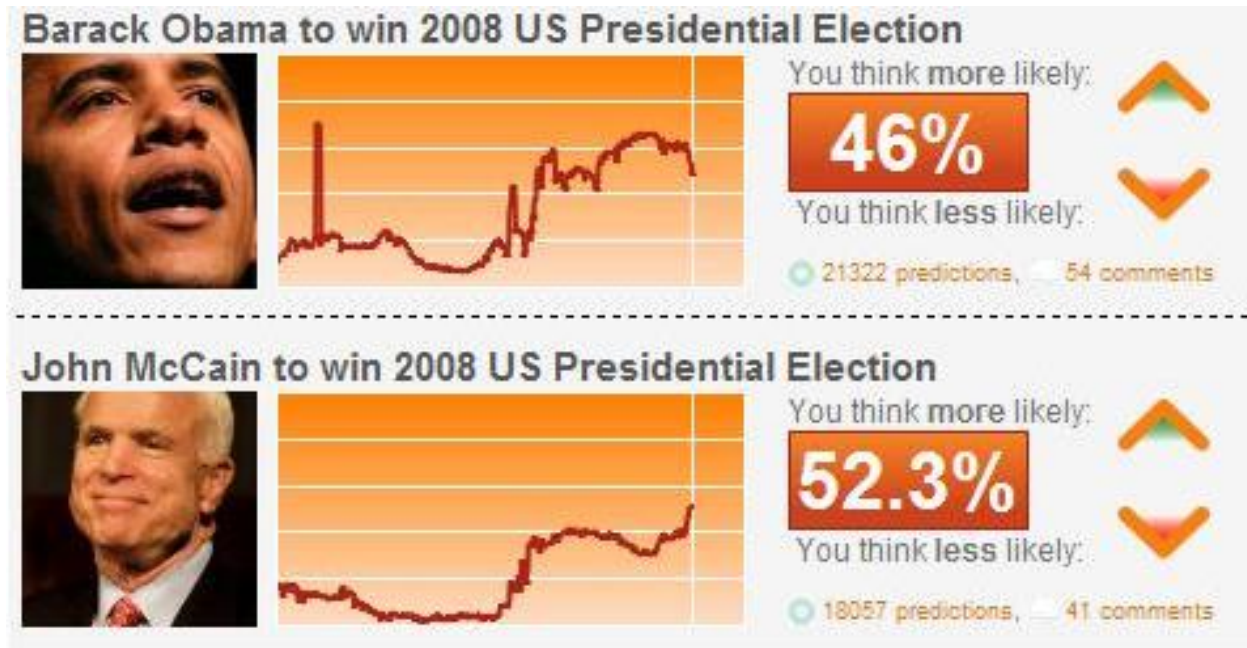


Kostnadsestimater



Prosjektets varighet

...og til mye annet rart



VÅRE ERFARINGER

Frode Torvund og jeg har undersøkt om prediksjonsmarked kan være et egnet estimeringsverktøy for små team i smidige prosjekter



Studie 1 : ”Feil funnet i systemtest”

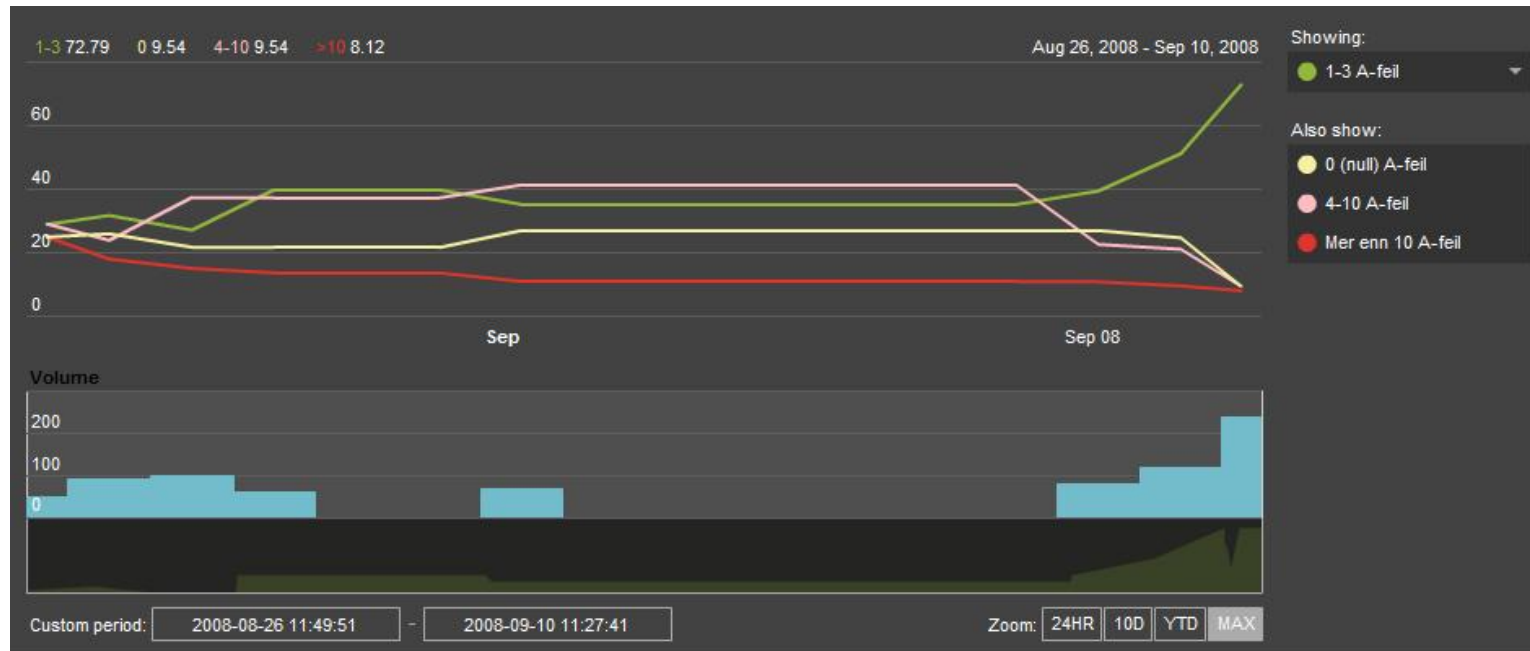
Prosjektet

- Del av større prosjekt, 7 deltakere
- Java- og web-teknologi
- Smidig utvikling med Scrum
- Ca 2 års varighet

Markedet

- ”Hvor mange A-feil blir funnet i systemtest”
- Fire svaralternativer (0, 1-3, 4-10, mer enn 10)
- Deltakere: prosjektleder og utviklere
- Varighet: 3 uker

Resultater



- 7 av 7 handlet i markedet
- Totalt 45 handler
- Handelen er konsentrert i kortere tidsintervaller

Studie 2 : "Status ved leveransedato"

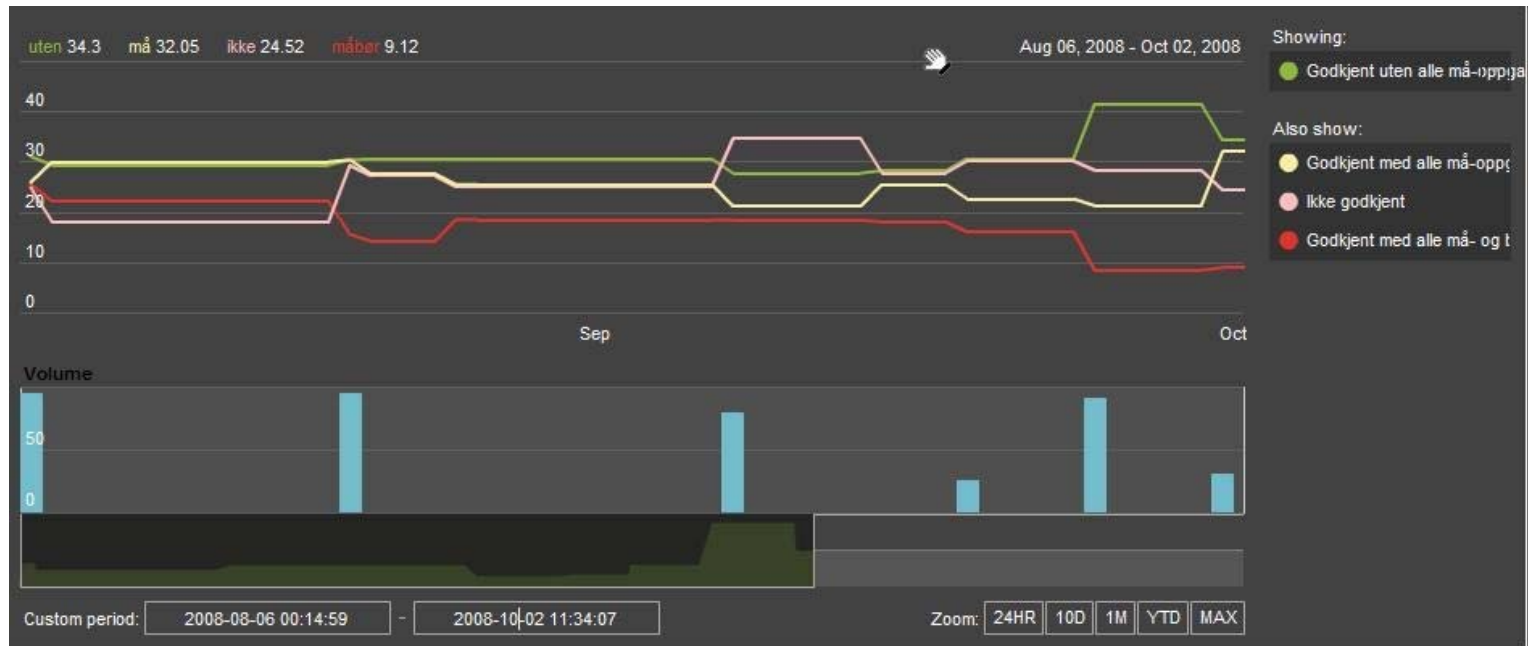
Prosjektet

- Timebasert prosjekt med 8 deltakere
- Ruby on Rails
- Smidig utvikling med Scrum
- Ca 6 måneders varighet

Markedet

- "Hva vil være status for leveranse 1 ved avtalt leveransedato?"
- Fire svaralternativer
- Deltakere: prosjektleder og utviklere
- Varighet: 8 uker

Resultater



- 8 av 8 handlet i markedet
- Totalt 52 handler
- Handelen er konsentrert i kortere tidsintervaller

LESSONS LEARNT

Vanskelig å evaluere estimeringsnøyaktigheten basert på studiene

De fleste forskningsstudier viser at prediksjonsmarkeder er minst like nøyaktige som ekspert-estimer

- Ingen klar støtte for dette i våre studier
- Illustrerte usikkerheten og dynamikken i prosjektene

Noen mulige problemområder:

- Mangelfull opplæring
- Lite egnet incentiv-modell
- Suboptimale marked og kontrakter
- Få deltakere
- Uegnet kontekst(?)

Det er nødvendig med grundig opplæring

Mail med instruksjoner var ikke tilstrekkelig opplæring

- Svært få kjente til prediksjonsmarkeder fra før
- Flere hadde problem med å forstå markedet
- Softwaren er ganske umoden og ikke selvforklarende

Flere forskningsstudier poengterer viktigheten av god opplæring

- Høye oppstartskostnader
- Uerfarne spiller vesentlig dårligere enn de med litt erfaring

Google rapporterer gode erfaringer med å bruke underholdningsmarkeder til opplæring (slik som "Hvem vinner fotball-VM?")



Incentiver er viktig, men det er ikke smart med premiering hvor "winner takes all" !

Vår incentiv-modell førte til uønsket atferd fra spillerne

- Noen mistet interessen når de ble liggende langt bak den som ledet
- Noen endret strategi når de lå langt bak (maksimere vinningsjansene, ikke mest rasjonelle alternativer)

De fleste forskningsstudier anbefaler at man har incentiver

- Resultatene er sprikende
- Men vær obs på at en stor premie kan påvirke prosjektgjennomføringen (bevisst eller ubevisst)

Loddtrekning virker å være et godt alternativ til "the winner takes all".

Hvor store skal incentiver være før det oppfattes som at det vil påvirke arbeidsinnsatsen?

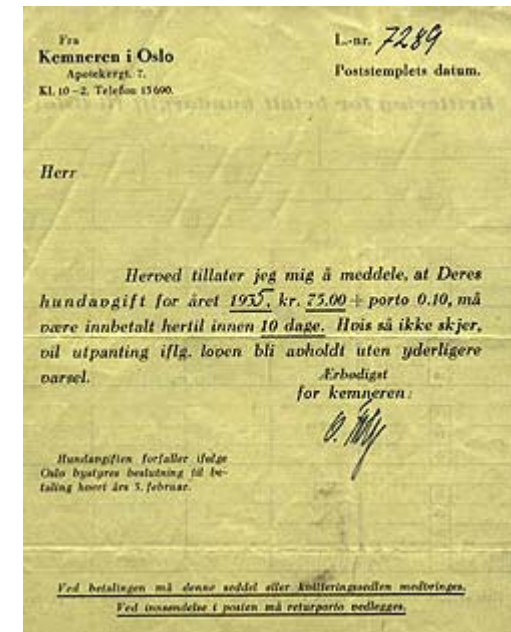
4%	Det vil uansett påvirke
26%	Gavekort på DVD
47%	Middag på restaurant
16%	Feriereise
7%	Det vil uansett ikke påvirke

Resultat fra spørreundersøkelse på JavaZone 2008

Det må purre!

I begge våre marked var handlene konsentrert i korte intervaller

- Markedet var sjeldent helt oppdatert, det kan gjøre det vanskelig å tolke resultater
- Øker graden av tilfeldighet
- Flere av deltakerne savnet jevnlig purringer



Studier rapporterer gode erfaringer med å kombinere handlingen med statusmøter i prosjektet.

Flere enn bare utviklerne bør delta i markedet

Liten variasjon blant deltakerne i våre markeder

- Alle deltakerne hadde tilgang til den samme informasjonen
- Lite politiske spenninger i prosjektene
- Transparent organisasjon

Forskningsstudier viser at gruppeestimeringsprosesser gjør det best når deltakerne har forskjellig bakgrunn

- Kunder, ledere, testere, sluttbrukere, styringsgruppe, osv
- Åpenhet kan være problematisk (internt og eksternt)
- Samlokalisering kan redusere verdien av prediksjonsmarkeder

I ditt nåværende prosjekt, tror du det ville vært mulig å få kunden til å delta i et prediksjonsmarked?

45%	Ja
55%	Nei

Resultat fra spørreundersøkelse på JavaZone 2008

Vær nøye med marked ("spørsmål") og kontrakter ("svaralternativ")

Markedet må være avgjørbart

- DÅRLIG: "Blir Age of Conan en suksess?"
- BEDRE: "Vil Age of Conan ha 400.000 betalende abonnenter i 01.01.09?"

Kontraktene må ikke overlappe

- DÅRLIG: [Uke 46], [46-50], [Uke 50 eller senere]
- BEDRE: [Innen uke 46], [uke 47-50], [etter uke 50]

Forskningsstudier viser at hvilke marked og kontrakter man velger kan ha stor påvirkning på estimeringsnøyaktigheten

- Svært usannsynlige hendelser blir typisk overvurdert
- Færre alternativer gir mer nøyaktige estimater
- Enkelt å manipulere deltakerne med ordbruk, og lignende

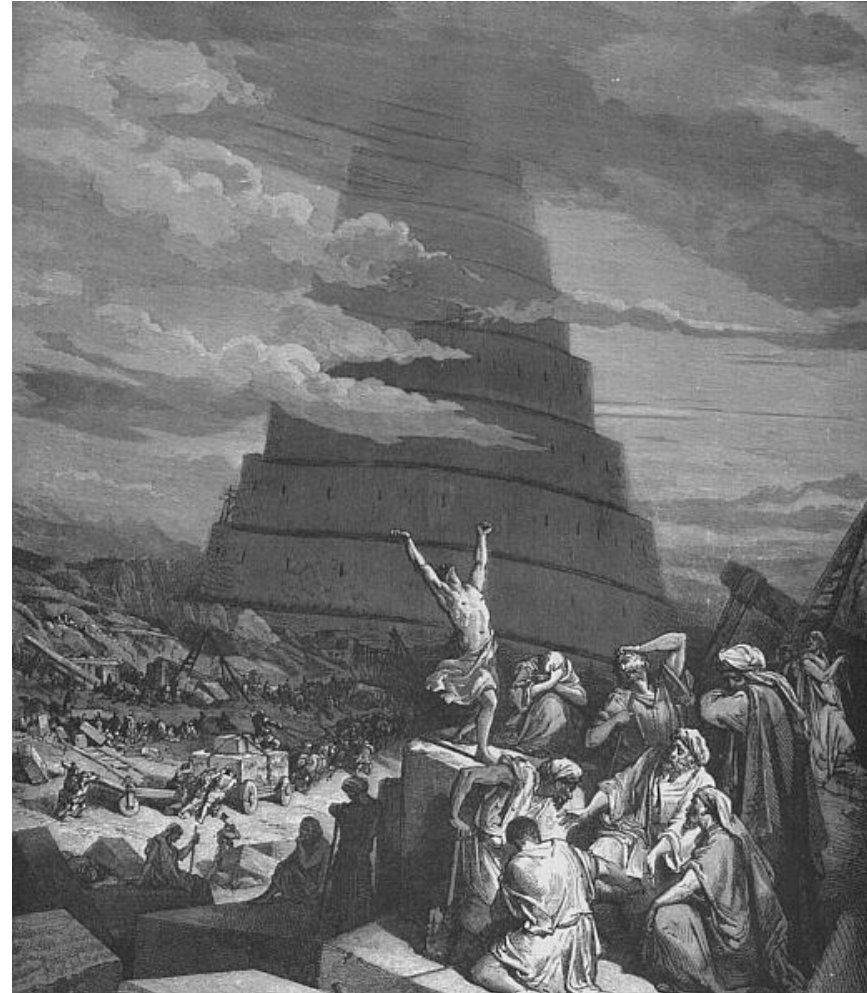
Positivt: Prediksjonsmarked medførte (mer) presis terminologi

Upresis terminologi er et av de store problemene innen estimering

- Kan hindre kommunikasjon og læring
- Ofte store forskjeller selv innen samme prosjekt

Stort fokus på å avklare terminologi i våre studier

- Eksempel: Skal feil som er feilaktig rapportert telle med?



Vi trenger flere studier av prediksjonsmarkeder i en software-utviklingskontekst

De fleste forskningsstudier omhandler domener som sports, politikk eller økonomi

- Flere hundre forskningsartikler
- Ikke mange studier i en softwareutviklingskontekst (noen få erfaringsrapporter)

Resultater er ikke nødvendigvis overførbare til softwareutvikling

- Scope endrer seg
- Mulig å påvirke resultatet
- Stor grad av usikkerhet
- Stor variasjon innen software-prosjekter
- Team-følelse

Positivt: Prediksjonsmarkder er gøy!

Deltakerne virket å like metoden

- Positive tilbakemeldinger i intervjuene
- Etablert fagprosjekt for å lage bedre software
- 60% av deltakerne i spørreundersøkelsen på JavaZone kunne tenke seg å ta i bruk metoden

Motivasjon er en viktig, men ofte glemt, suksessfaktor innen prosessforbedring



Noen nyttige referanser

Webressurser

International Journal of prediction markets

Prediction markets SIG at www.forecastingprinciples.com

Bok

"The Wisdom of Crowds" av James Surowiecki

Artikler, rapporter, og lignende

"Hvordan og hvorfor gjøre beslutninger med prediksjonsmarked?"
av Øyvind Rogne

"The promise of prediction markets: A round-table"
av Renée Dye

"Prediction Markets"
av Justin Wolfers & Eric Zitzewitz.

Q & A ?