

Transforming Business with AI

Guest lecture at NHH

Erling Risa
Sparebanken Vest

September 24, 2024

Briefly About Me

- PhD from NHH in 2019
- At Sparebanken Vest since then
- Solving business problems with statistics and machine learning / AI

Learning Outcomes

Upon completion of the course, the student can:

- Identify applications and implications of AI in business and research
- Recognize the ethical considerations, privacy concerns, and risks associated with AI technology

Outline

Generative AI

- Various use cases

“Traditional” AI

- Anti money laundering case

Generative AI

Productivity

- Writing
- Some coding tasks
 - Scaffolding
 - Debugging
 - Testing
 - Documentation
 - Code completion
- Giving the models access to internal data makes them a lot more useful
 - This has major implications for data governance
 - GDPR

Summarizing

- Banks face a lot of regulation
- It would be nice if AI could help us keep track of it all
- What happens if we give an LLM access to regulatory documents and ask it to summarize them?

Pretty good...

What does Finanstilsynet say about PD-modeling?



Finanstilsynet, in its **Rundskriv 3/2021**, provides detailed guidelines for Probability of Default (PD) modeling as part of the Internal Ratings-Based (IRB) approach used by banks for credit risk assessment. The main aspects covered include:

1. Long-Term Default Rate:

- PD estimates must reflect long-term average default rates, in accordance with Article 180 of the Capital Requirements Regulation (CRR). The data should include not only good economic periods but also **downturn conditions** to ensure that the models are robust under different market conditions.

2. Incorporating Economic Downturns:

- A minimum of **20% of the data** used in PD models should come from periods reflecting an economic downturn. Finanstilsynet points to the **bank crisis of the early 1990s** as the last severe downturn in Norwegian history. Where banks lack sufficient downturn data, they must estimate default rates for downturn periods that are representative of their portfolios.

3. Safety Margins:

- PD estimates must include **safety margins** to account for uncertainty in data and model deficiencies. These margins help address limitations such as a lack of downturn data or weaknesses in historical information. For example, portfolios with minimal historical defaults, such as real estate loans, require specific safety margins.

4. Data Requirements:

- The data used for PD estimation must be **relevant and sufficient** to capture both good and bad economic times. Banks are required to use data covering at least the last five years, though longer data histories are encouraged, especially when modeling for more cyclical or sensitive portfolios.

5. Model Validation:

- PD models must be validated regularly by comparing estimated default rates with observed defaults. This validation should focus both on the **ranking ability** of the model (how well it differentiates risk) and the **absolute level of PD estimates** compared to real outcomes. Validation must occur at least annually.

In summary, Finanstilsynet sets a high bar for PD modeling, requiring comprehensive data that includes downturns, safety margins for uncertainties, and rigorous validation to ensure the accuracy and robustness of PD estimates.

...but not perfect

Original:

For lån med pantesikkerhet i bolig er det særskilt usikkerhet knyttet til datakvaliteten fra bankkrisen. Usikkerheten forsterkes av at bankens øvrige datagrunnlag er preget av svært gode år i norsk økonomi og høy boligprisvekst. Banken må forutsette en misligholdsrate på minst 3,5 prosent ~~en alvorlig nedgangskonjunktur~~. Nivået er beregnet ut fra observerte problemlånsandeler under bankkrisen justert blant annet for forholdet mellom beholdningen av problemlån og andel nye mislighold. For å ta hensyn til svakt datagrunnlag i segmenter med lite mislighold må banken tillegge sikkerhetsmargin på særskilt lave misligholdssannsynligheter, slik at alle boliglån har en PD på minst 0,2 prosent.

ChatGPT:

- For lån med pantesikkerhet i bolig, er det satt en forutsetning om en misligholdsrate på minst 35% i en alvorlig nedgangskonjunktur. For boliglån med lav misligholdssannsynlighet, kreves det en minimum PD på 0,2% for å reflektere usikkerhet i datagrunnlaget.





















Chatbot

Note: This example is a few months old, and this is a rapidly changing area

- Proof of concept
- Build a searchable index of the bank's webpage and put a chatbot on top of it
 - Retrieval-augmented generation (RAG)

Building the Index

- Build the index in Azure AI Search
- This gives us a number of searchable files in a storage container
 - These will form the chatbot's knowledge corpus

	baerekraft.html
	bankid.html
	bedrift.html
	bli-kunde.html
	boliglanskalkulator.html
	dagligbank.html
	eiendom.html
	forsikring.html
	index.html
	index1.html
	index2.html
	kort.html
	kundeservice.html
	kundeutbytte.html
	laaneforberedelse.html
	lan.html
	sikkerhet.html
	spare.html
	un.html
	utenlandsbetaling.html

Searching the Index

 sparing 

Search

Results

```
1  {
2    "@odata.context": "https://spv-laeredag-search.search.windows.net/indexes('spv-er')/$metadata#docs(*)",
3    "@search.answers": [],
4    "value": [
5      {
6        "@search.score": 2.4562302,
7        "@search.rerankerScore": 2.697150707244873,
8        "@search.captions": [
9          {
10             "text": "Grønn barnesparing Sparing i bærekraftige fond er bra for barna og planeten de skal arve. Sta
11             "highlights": "Grønn<em> barnesparing Sparing</em> i bærekraftige fond er bra for barna og planeten de
12           }
13         ],
14         "content": "Hopp til hovedinnholdet\n\nBli kunde\nLogg inn\n\tLogg inn privat\n\t\n\tLogg inn bedrift\n\nMe
15         "filepath": "spare.html",
16         "title": "Hopp til hovedinnholdet\n\n\n\n\nBli kunde\nLogg inn\n\tLogg inn privat\n\t\n\tLogg inn bedrift
17         "url": "https://laeredagstorageer.blob.core.windows.net/webpage-spv-er/spare.html",
18         "id": "aHR0cHM6Ly9sYWVyeZWZ3N0b3JhZ2Vlci5ibG9iLmNvcuUud2luZG93cy5uZXQvc3B2LWVyeLWNoZW5rcy9hSFwiY0hNNkx5OXNl
19         "chunk_id": "0",
20         "last_updated": "20240424055810"
21       },
22     ]
23   }
```

Querying the Chatbot I

Hva slags sparing tilbyr dere?

Sparebanken Vest tilbyr flere forskjellige sparealternativer, inkludert kontosparing, fondssparing, barnesparing, pensjonssparing og BSU-sparing ^{1 2}. Du kan velge den sparemetoden som passer best for deg og dine behov. Hvis du ønsker mer informasjon om de forskjellige sparealternativene, kan du besøke bankens nettside eller ta kontakt med en av bankens rådgivere ².

2 references ▾

1 kundeutbytte.html - Part 1

2 spare.html - Part 1

AI-generated content may be incorrect

- Reasonable answer
- References

Querying the Chatbot II

Hvor mange kontorer har dere?

Sparebanken Vest har flere kontorer på Vestlandet ¹. Du kan finne ditt nærmeste kontor ved å besøke bankens nettside og bruke kontofinneren ² ¹. Jeg beklager, men jeg har ikke informasjon om det eksakte antallet kontorer.

2 references >

AI-generated content may be incorrect

- Does not hallucinate
- Unexpected that it does not have the answer, as the information is on the website
 - It turned out that the indexing procedure had failed to index parts of the website
 - The bot is only as good as the data it has available

Bergen > Jonsvoll, Bergen sentrum > Fyllingsdalen > Indre Årna > Lagunen > Åsane > Vestkanten	Hardanger og Voss > Voss > Kinsarvik > Norheimsund > Odda	Møre og Romsdal > Volda > Ålesund
	Nordhordland	Sunnhordland > Bremnes, Bømlo > Leirvik, Stord

Chatbot Conclusion

- Building a chatbot like this is a relatively easy way of getting started with generative AI
- Potential uses both internally and externally
- The knowledge corpus is key
- Proper configuration is also essential

Working with CRM Data

- Customer relationship management has many important uses
 - Not always easy to work with
- LLMs are highly capable of transcribing audio
 - We can use this to transcribe customer calls
- Textual data from customer contact can be further summarized and classified

“Traditional” AI

Motivation

- Present a “mature” case
 - Common issues
- “Traditional” AI is (still) often the right tool

Anti Money Laundering

- Banks need to have high quality anti money laundering (AML) systems in place
 - Know your customer (KYC)
 - Monitor for suspicious activity
- In cases of suspected money laundering, we are required to:
 - Report to the authorities
 - Put appropriate measures in place

Problem Definition

- Regular portfolio screening
 - How do we prioritize which customers to look at?

The Challenge

- This is a complicated task for several reasons
 - Large number of customers
 - Many features of the customers that are relevant
 - Large volumes of transactions
- It is also a difficult phenomenon to observe
 - Money launderers are actively trying to conceal their activities
 - Money laundering schemes change over time

Approaches to AML

- Rule-based
 - Define a set of criteria that are indicative of money laundering
 - Clear what is going on (at least at first)
 - Difficult to specify and maintain high quality rules
- Model based
 - Use an algorithm to identify high risk customers
 - More flexible
 - Potentially complicated
- There is a place for both

Supervised or unsupervised learning?

Supervised:

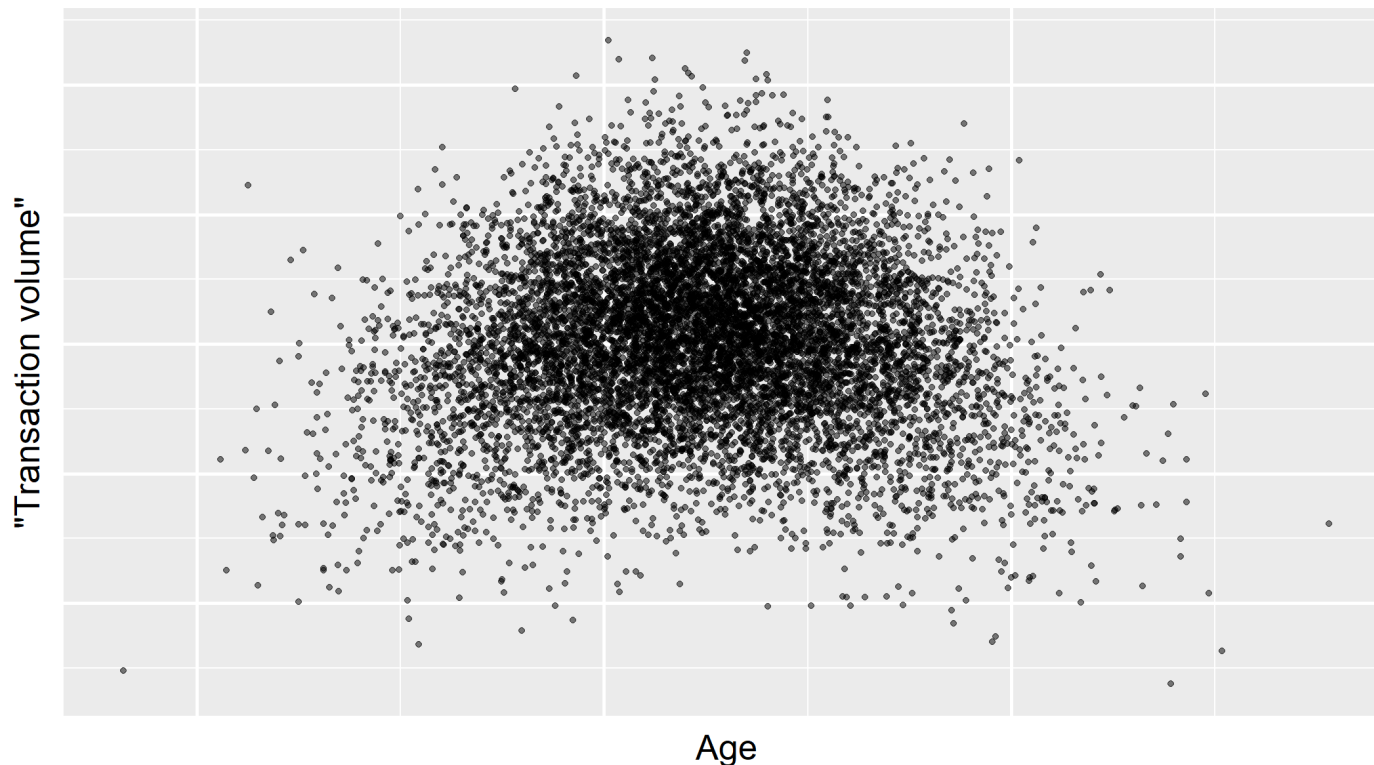
- Use known examples of known money laundering to teach the model what to look for

Unsupervised:

- Methods that do not rely on labelled data

Unsupervised Learning

- Let us simulate some data to illustrate how unsupervised learning can be used for anomaly detection



Anomaly Detection



Anomaly Detection

- Here we used an isolation forest to make an anomaly score
 - A measure of how “unusual” an observation is
- This approach has a lot going for it
 - No need for labelled data
 - Maybe this is the only feasible option
- However:
 - All deviations are treated equally
 - Unusual \neq money laundering

Supervised learning

- We decided that we had enough data to enable a machine learning algorithm to learn from past examples of money laundering
- Such a model will target exactly the behavior we have reported in the past
- Risk “blind spots” for money laundering schemes that are rare or new
 - Good reason **not** to have this model be the only control mechanism

How did it turn out?

- The new model provided a very useful supplement to the existing process
 - Significant share of reported cases
- Far better at detecting complicated patterns
- Lower false positive rate
- Easier to adapt to changes
 - But making changes to a complex model requires care

Interpretability

- AI models are often referred to as “black boxes” because they are difficult to interpret
- Initially we had no way of explaining to end users why a customer received the score they did
- This made the handling of flagged customers inefficient

Explainable AI

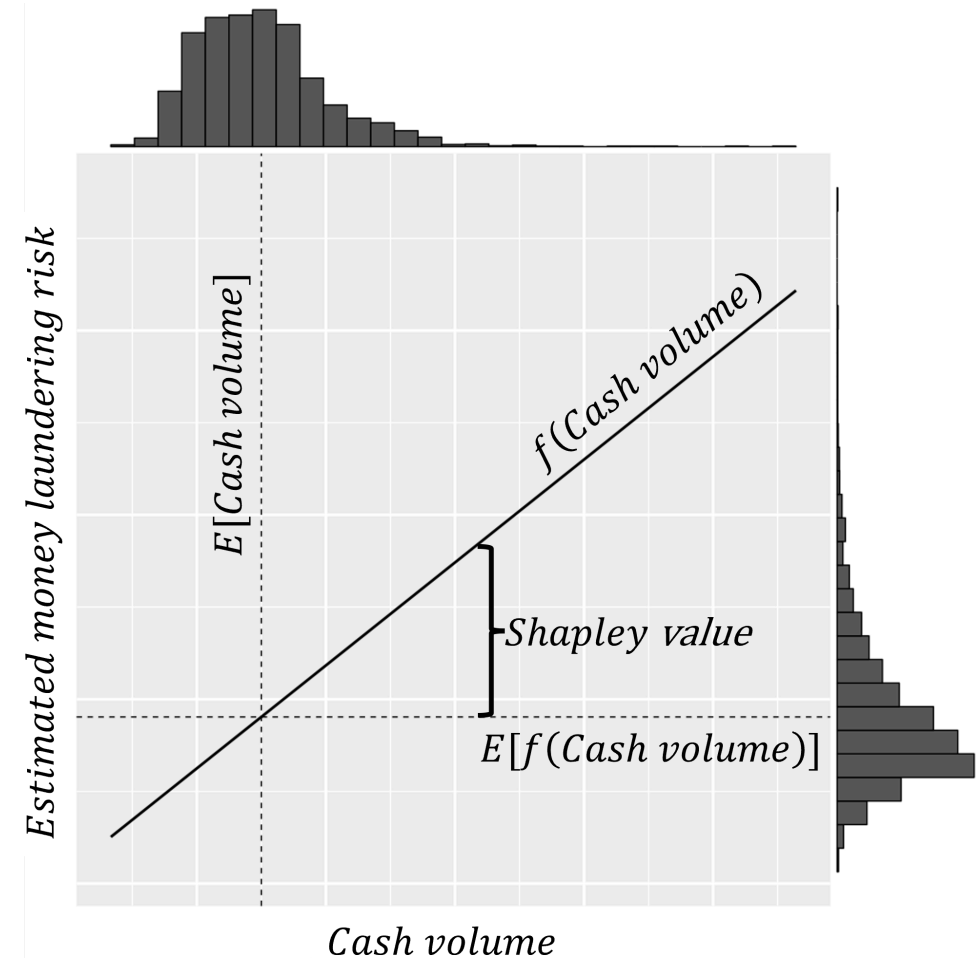
ChatGPT:

Explainable AI (XAI) refers to artificial intelligence systems designed to provide clear, understandable explanations for their decisions, predictions, or actions.

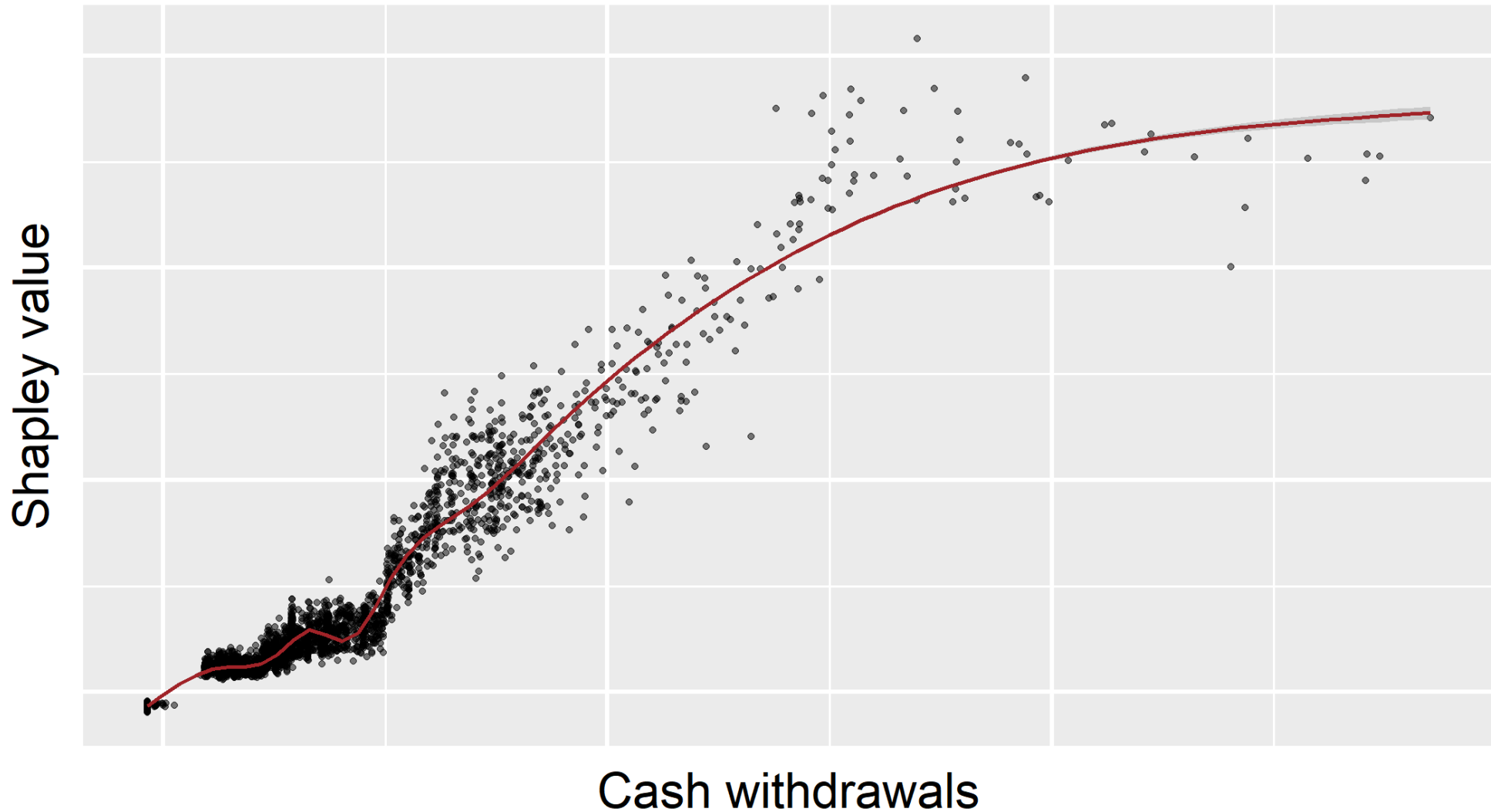
- For LLMs:
 - Which nodes are being activated?
- For “traditional” machine learning algorithms:
 - Which variables have the most influence on the prediction (and in what way)?

Shapley Values

- Shapley values are a way of attributing the prediction of a model to the different features
- Intuitively, it is the average marginal contribution of a feature to the prediction
- Intractable to calculate exactly for complex models, but can be approximated



Shapley Values in Practice



Problem solved?

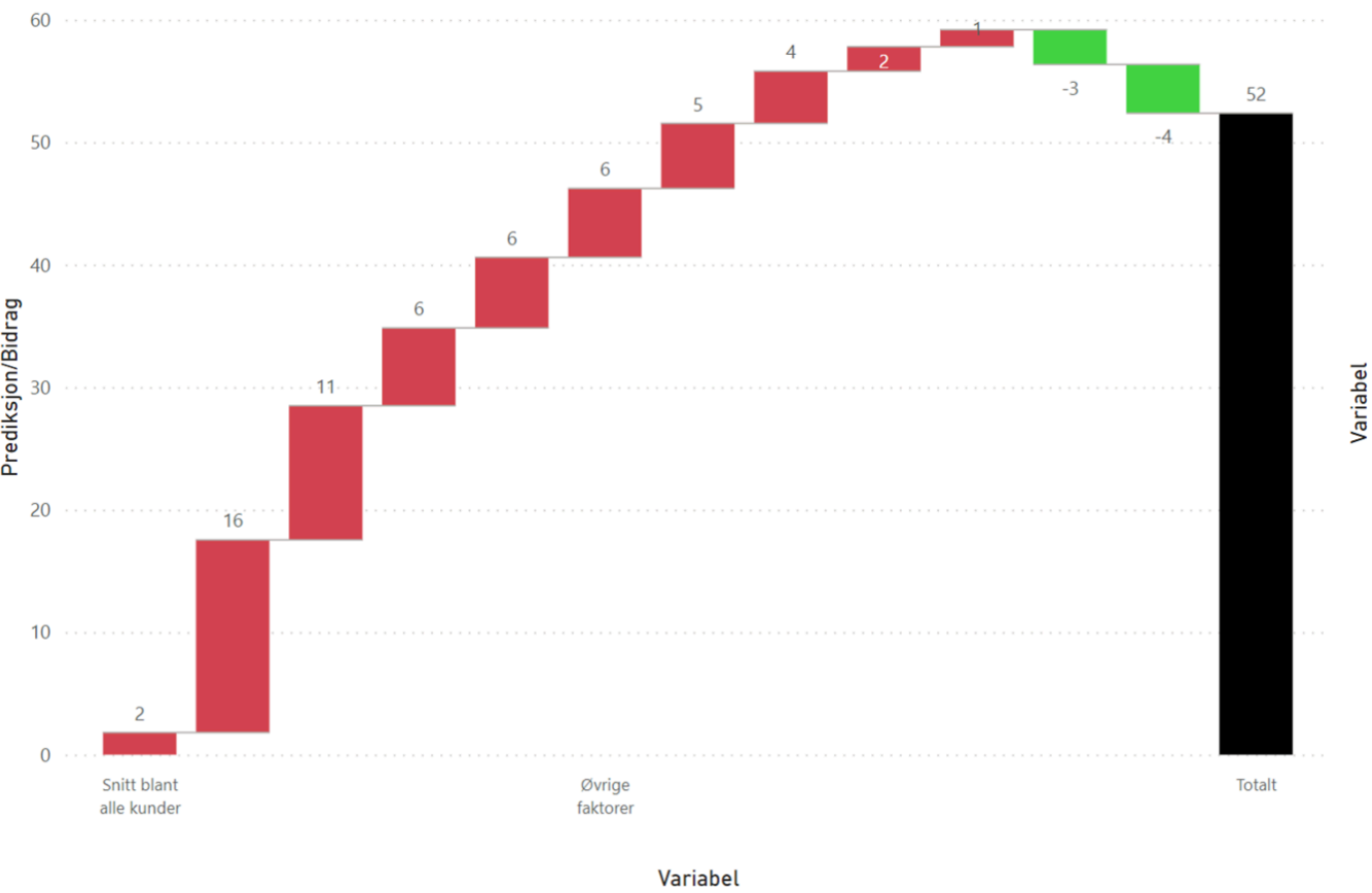
- How do we serve up-to-date Shapley values to users on demand?
 - Important to keep in mind when designing the process
- Privacy concerns
 - Justifiable purpose?
 - Access control
 - Logging
 - Data minimization

Dashboard

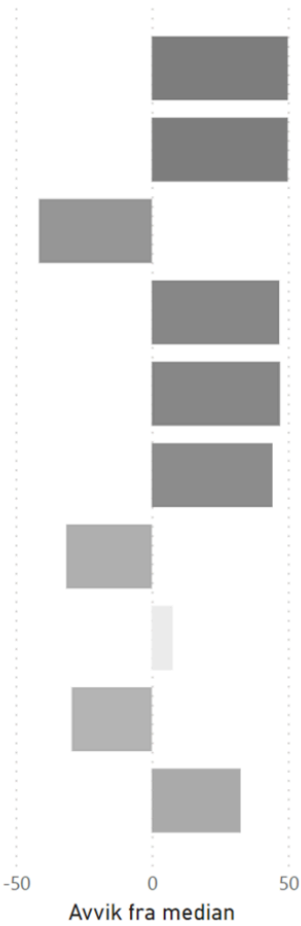


Bakgrunn for AHV-score

Øk Reduser Totalt



Kundeprofil relativt til øvrig kundemasse

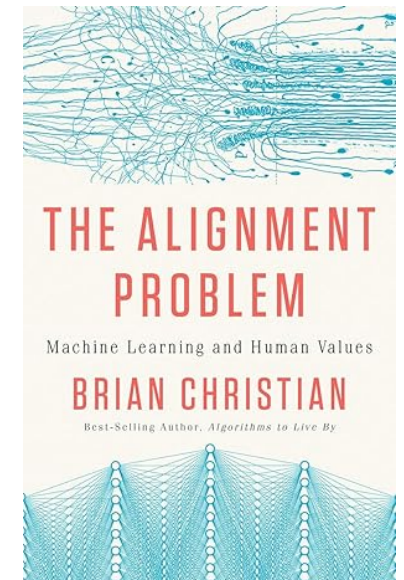


Alignment

ChatGPT:

The process of ensuring that an AI system's goals, behaviors, and actions are consistent with human intentions, values, and ethical standards to prevent unintended or harmful outcomes.

Or: “Is the model behaving as intended?”



Why is alignment difficult?

- It is difficult to specify exactly what we want the model to do
 - Especially true for LLMs
- And we cannot observe in detail how the model arrives at conclusions

Unintended Consequences



- When we first implemented the model, Bulder was small, but growing fast
- Some Bulder customers exhibited a transaction pattern that was prone to being flagged by the model
 - Unintended and unanticipated
- Because we had “a human in the loop”, this did not affect any customers

Conclusion

- AI is incredibly powerful
- But:
 - These models are complex
 - We do not fully understand how they work
 - We are unable to specify exactly what we want
- Weigh costs against benefits