PANEL: Analyzing AI Advancements & Applications for Fighting Financial Crime & Improving Compliance

PANEL: Analyzing AI Advancements & Applications for Fighting Financial Crime & Improving Compliance

MODERATOR

Changxin Miao - Machine Learning Engineer - **ING**

PANELIST

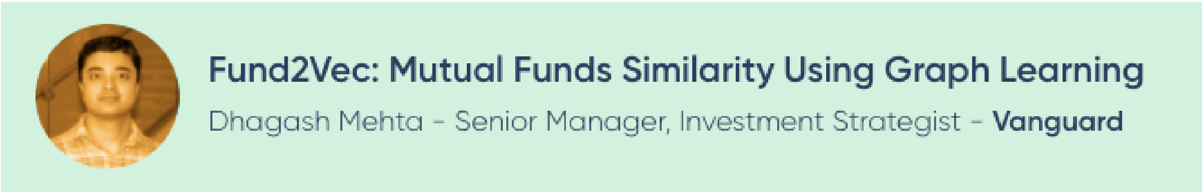
Mehrdad Mamaghani - Head of Department, Data Science & Applied AI, Anti-Financial Crime - **Swedbank**

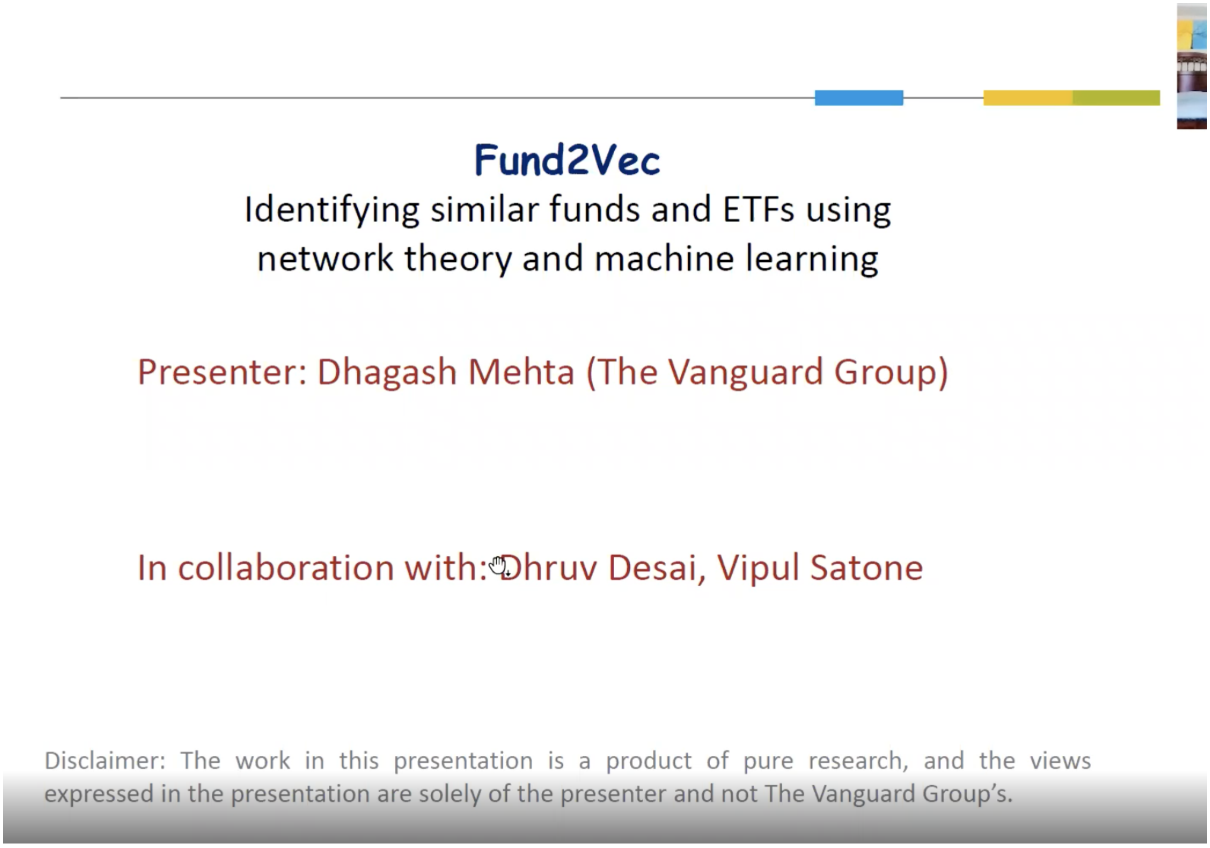
PANELIST

Chris Merz - Vice President Security and Decision Products - **Mastercard**

- Swedbank works with only one cloud solution (Azure)
- Mastercard works with several cloud solutions (AWS, GCP and Azure), and also, implements their own tools in cloud
- Mehrdad mentioned the use of graph NN for crime prevention modeling
- Chris mentioned the use of ensemble algorithms for classification, also mentioned the challenge of providing explanation on rejections

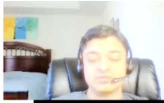
Fund2Vec: Mutual Funds Similarity Using Graph Learning





1.40

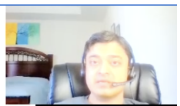
Introduction: Product Similarity



- *Question 1*: Given a product, what are other similar products?
- *Question 2*: How similar are the given two products? (Or, different from each other?)
- 'Product Similarity' is also known as 'Peer Analysis', 'Competitors Analysis', etc.
- A very frequently arising problem in most business areas.
- In our case, the problem may specialize to finding similar funds or ETFs.

1.40

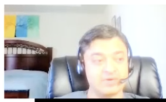
Introduction: Why Product Similarity?



Multiple applications:

1. Sales and Marketing: knowing a customer has a competitor's fund, proactively convince the customer to switch to a similar home-grown product;
2. Alternative Portfolio Construction: For a given portfolio of funds consisting of competitors' funds, construct an alternative portfolio with the same risk-return profile but consists of only home-grown funds.
3. Portfolio Diversification: two or more similar funds in a portfolio may unintentionaly reduce diversification;
4. Similar fund with different theme: e.g., find similar fund but with other attribute (e.g., similar ESG fund)
5. Competitors' Analysis: A fund manager can compare various aspects of their managed fund with other similar funds managed by competitors;
6. Tax Loss Harvesting: Move from one fund to another similar one for tax-loss harvesting;
7. Launching New Products: Launch a new fund similar to one popular in specific markets.

1.40



Introduction: Current Approaches for Product Similarity?

1. Third-party Categorization: e.g, Morningstar/Lipper categories. (Known to partly rely on qualitative approach, partially a black-box and sometimes irreproducible process. More importantly, no ranking)

2. Compute the overlap between two portfolios (with the Jaccard index, weighted Jaccard index, etc.) (captures the bigger picture but need to be careful if granular details are needed)

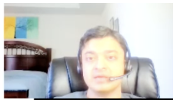
3. Compute the Euclidean distance between pairs of portfolios in the chosen variables-space (captures linear relationship)

4. Compute the cosine similarities between vectors corresponding to different portfolios in the chosen variables-space (captures linear relationship)

5. Many other unsupervised machine learning techniques such as unsupervised clustering (usually linear relationship, or doesn't scale well)

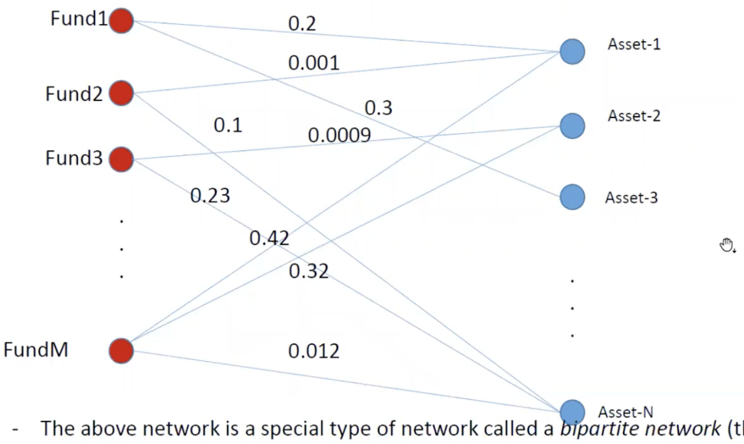
Our idea: reformulate the data of mutual funds and assets as a network, use a graph neural network to identify the embedded representation of the data, and compute similarity in the learned lower dimensional representation.

1.40



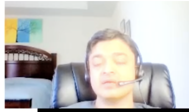
Product Similarity: Funds-Assets Network

Our idea is to represent the funds and assets as a network (Delpini et al 2018):



- The above network is a special type of network called a *bipartite network* (there are two distinct types of nodes here: funds+ETFs and assets).
- Bipartite networks are used in investigation of various social networks, movie-actors networks, protein-protein interactions, genome networks, flavor-ingredient networks, etc.

Product Similarity: What is similarity on a network?



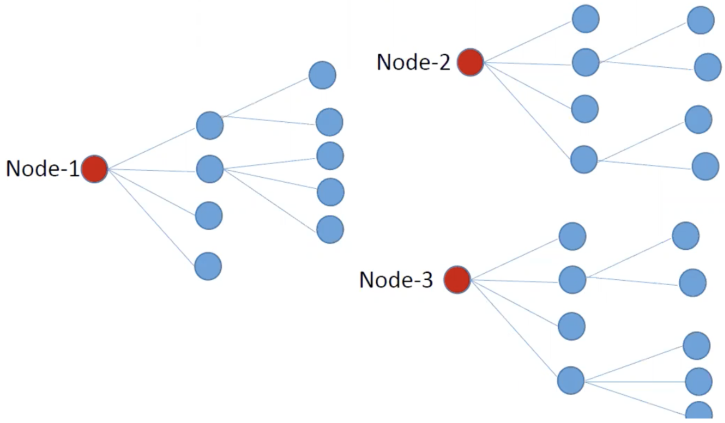
- The problem of finding similar products is now transformed to finding similar nodes on the network.

- There are many ways nodes can be similar to one another on a network.

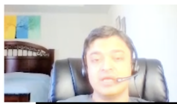
E.g.,

- Nearest-neighbor node degree similarity:

Node-1 is more similar to Node-3 than Node-2



Product Similarity: What is similarity on a network?

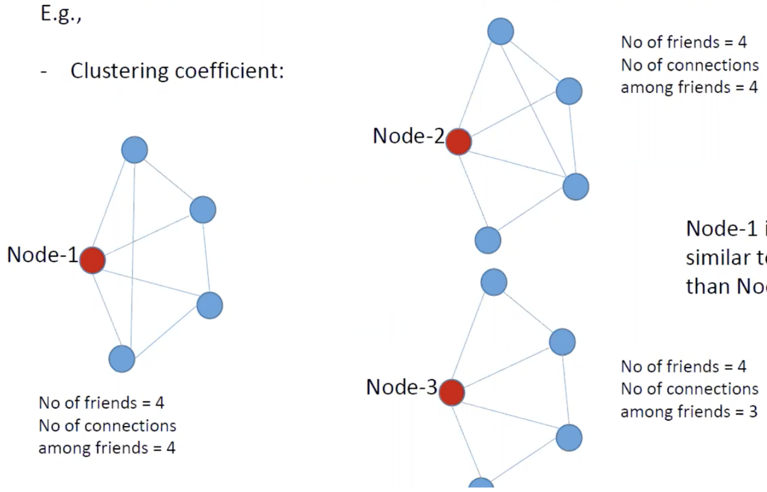


- The problem of finding similar products is now transformed to finding similar nodes on the network.

- There are many ways nodes can be similar to one another on a network.

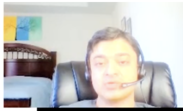
E.g.,

- Clustering coefficient:



Node-1 is more similar to Node-2 than Node-3.

Product Similarity: What is similarity on a network?

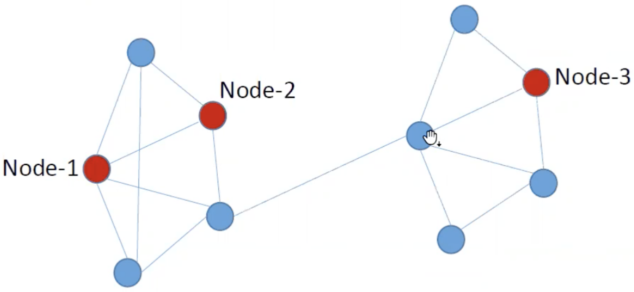


- The problem of finding similar products is now transformed to finding similar nodes on the network.

- There are many ways nodes can be similar to one another on a network.

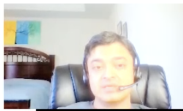
E.g.,

- Homophily:



Node-1 is more similar to Node-2 than Node-3 because the first two are in the same 'community'.

Product Similarity: What is similarity on a network?



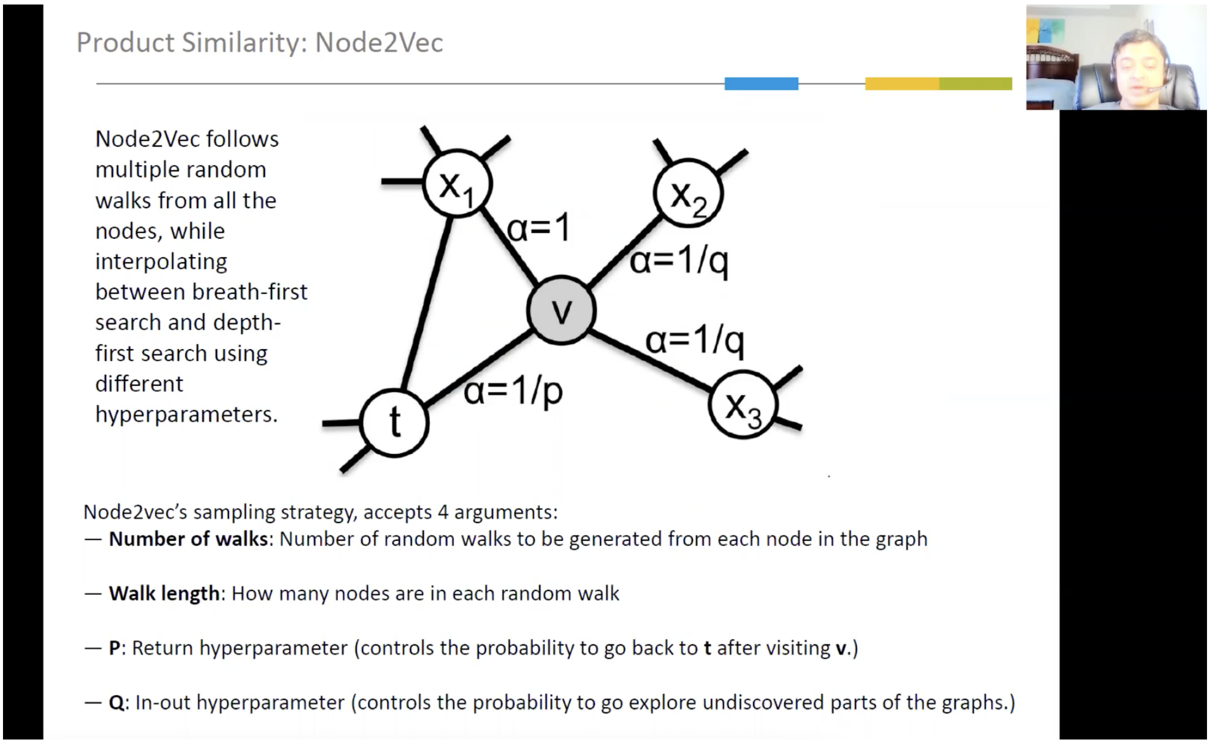
- The problem of finding similar products is now transformed to finding similar nodes on the network.

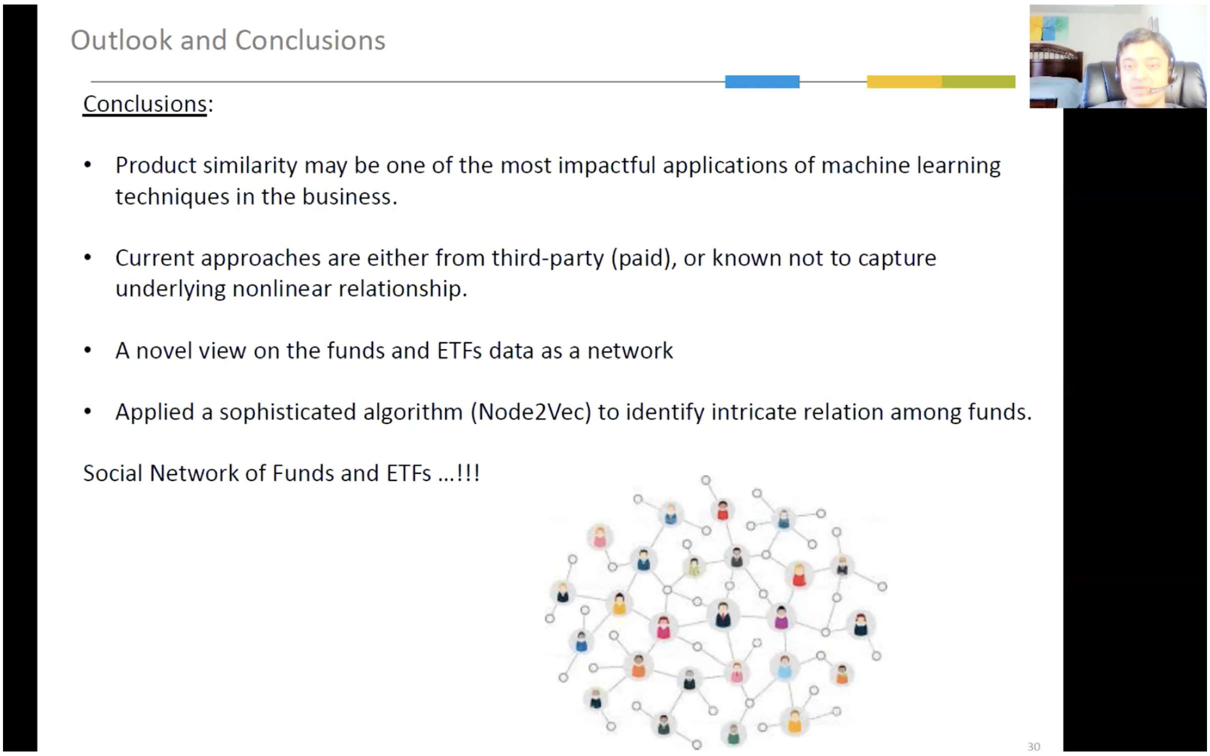
- There are many ways nodes can be similar to one another on a network.

- Idea: create as many such network similarity features as possible, and do unsupervised learning to identify overall similar nodes.

- Have to create many such features? Computing them for all the nodes may not even be scalable.

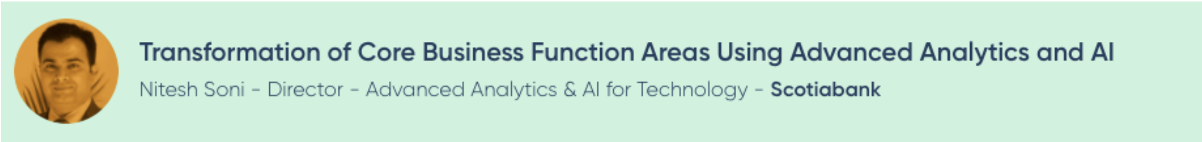
- New idea: Node2Vec (2016). Learn the features (representation learning) from the raw data.

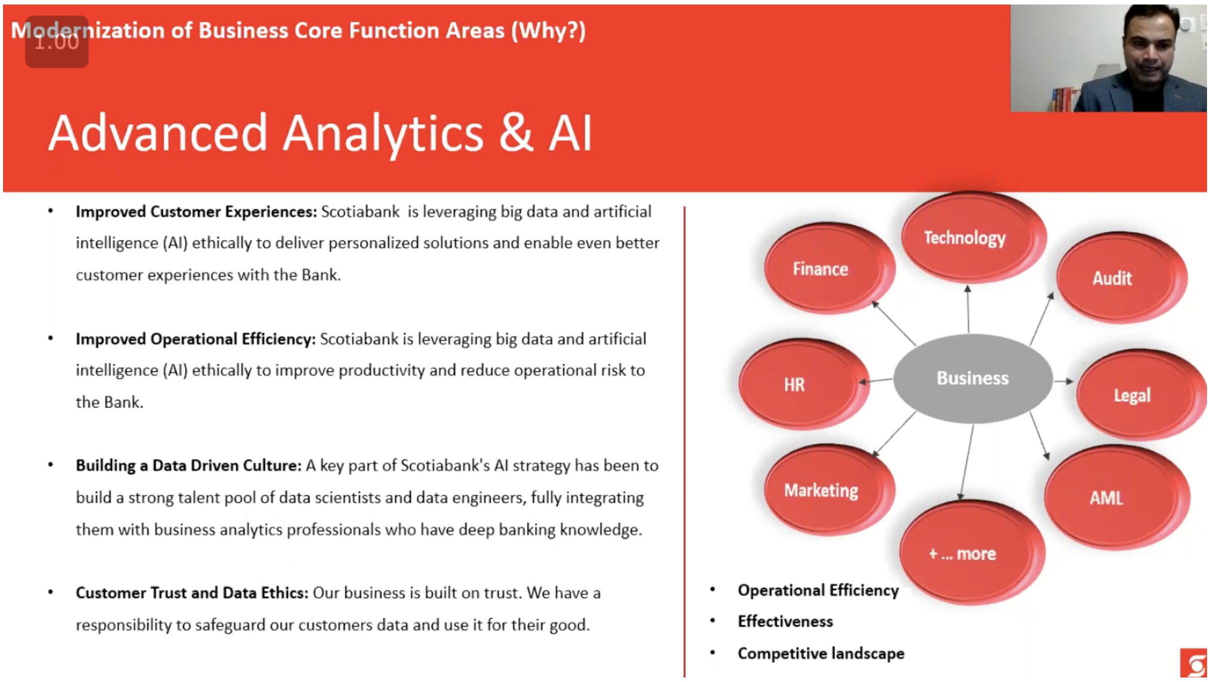


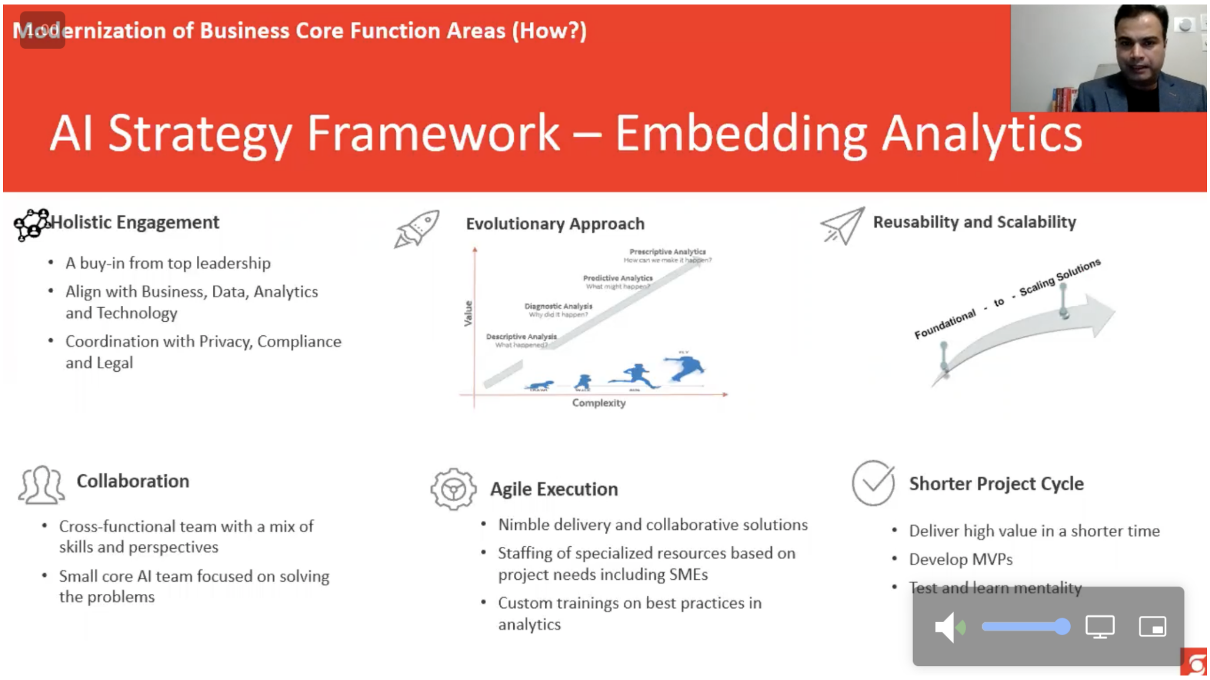


- Although he mentioned several motivations, he later focused more on the finding similar funds with different themes
- More technical perspective, he elaborated the project into a network problem
- I'm a big fan of Vanguard and loved to see a bit more about it :)

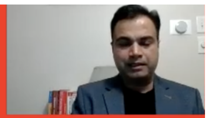
Transformation of Core Business Function Areas Using Advanced Analytics and AI





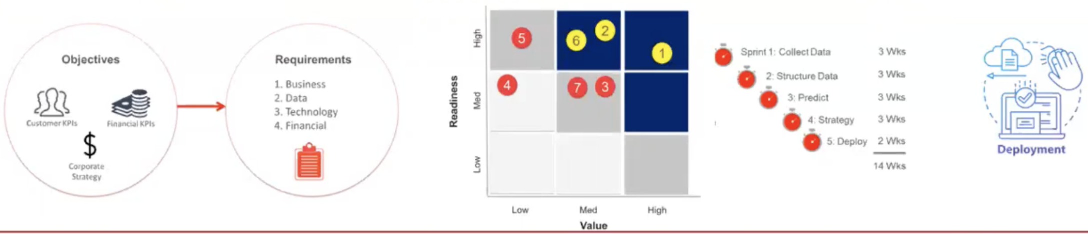


An example of framework



AI Engagement and Delivery Framework

ENGAGEMENT FRAMEWORK				DELIVERY FRAMEWORK	
INITIATION	DISCOVERY	ESTIMATES	PLANNING	EXECUTION	POST-RELEASE



- What are stakeholders pain-points?
- How Analytics can help to Solve them?
- What are the main requirements?

- What is the business value?
- What is the readiness of data, technology, people etc.

- Agile Execution

- Deployment
- Continuous feedback (automated or non-automated way)

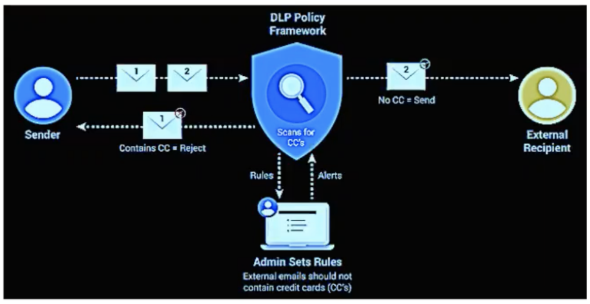
Use Case Example – Staying Vigilant and protecting bank from Cyber attacks

Cyber Security – Data Loss Prevention

- **Data Loss** can happen in multiple ways – **theft or loss of laptops, via emails (unknowingly or inside malicious activity), connecting to external network** etc. etc.

- **Embedding AI to detect and prevent Data Loss –** by enhancing traditional policy based rule and analysis massive alerts data

- Analyzing the **massive volume of email logs**
- **Pattern Analytics**
  - Frequency of corporate email to corporate emails, corporate email to personal email etc.
  - Email with or without attachment
  - Timing of email etc.
- **Text Mining using NLP**
  - Content of emails (e.g. contains cc or other sensitive information)



- **Reduction in False Positive Alerts using ML and Deep Learning**
  - Integrate DLP with user and entity behavior analytics to prioritize the avalanche of alerts
  - Focus on the legitimate alerts

Transforming Facebook Finance Operations with AI

Transforming Facebook Finance Operations with AI

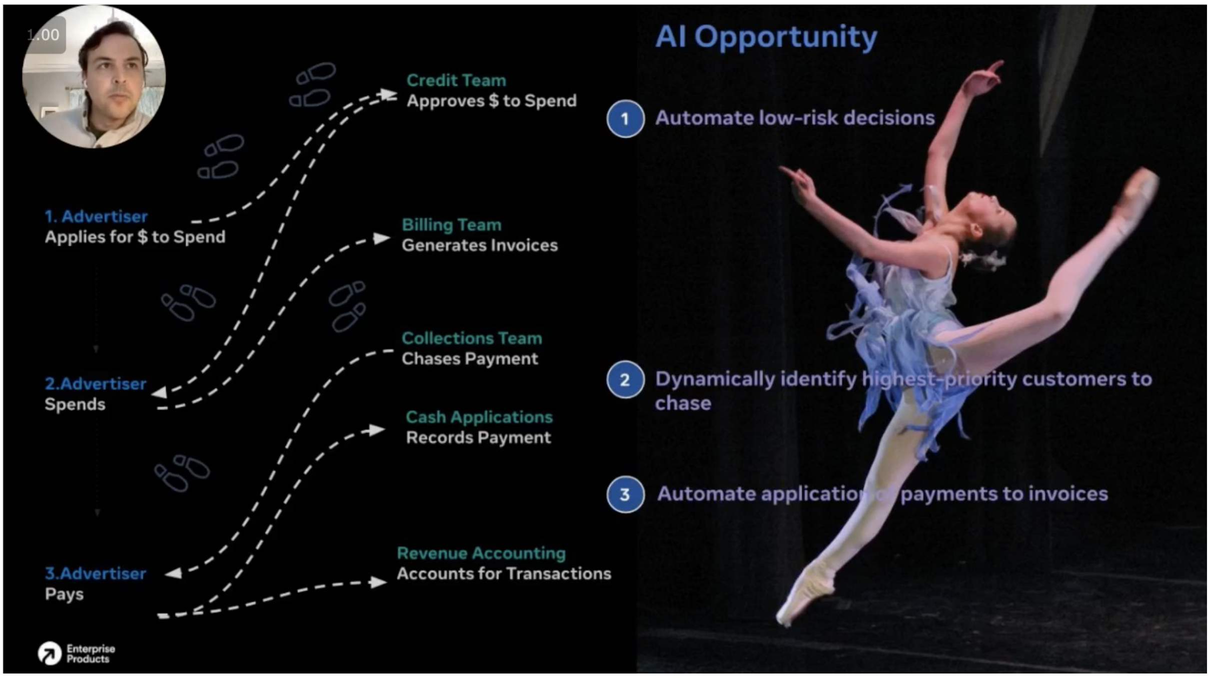
Julie Drew - Applied Research Manager, Enterprise Products - **Facebook**

Transforming Facebook Finance Operations with AI

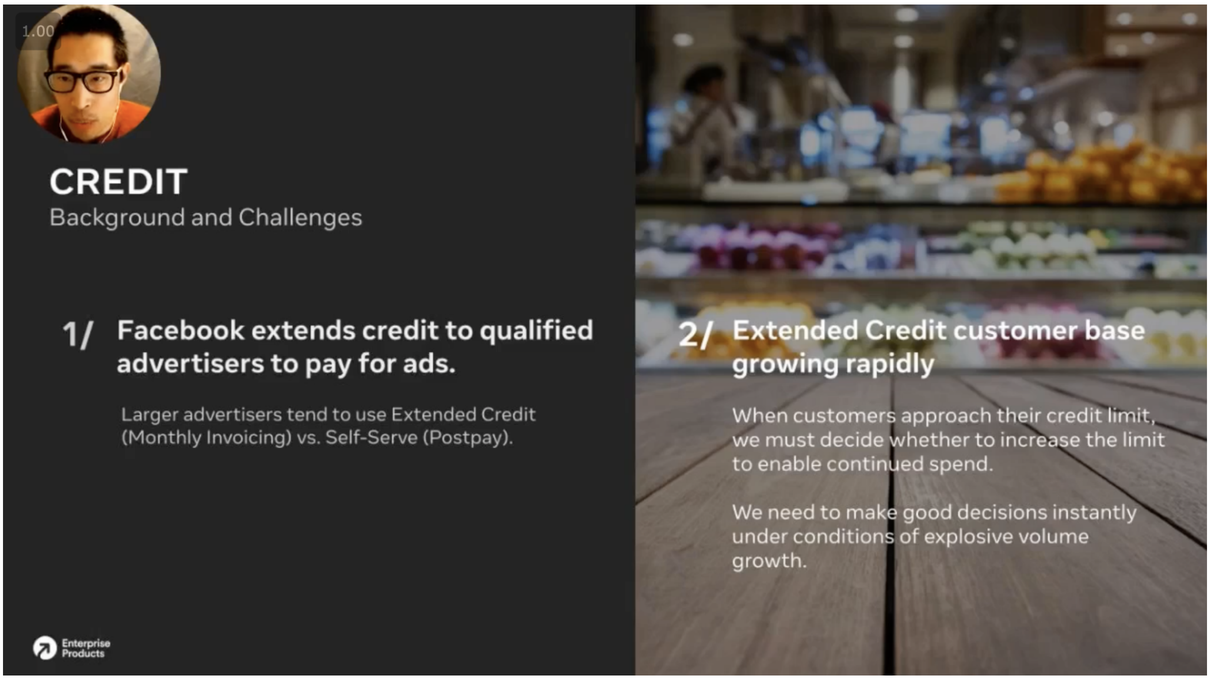
Paul Brandenburg - Program and Analytics Manager - **Facebook**

Transforming Facebook Finance Operations with AI

David Chi - Applied Research Scientist - **Facebook**



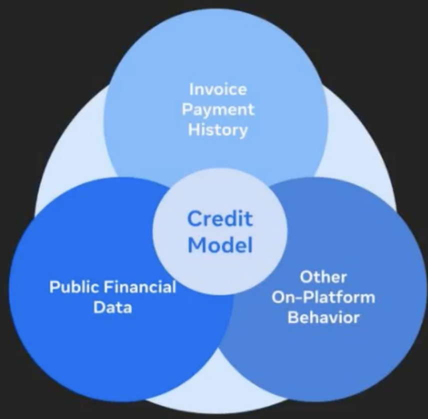

1.00
Enterprise Products Applied Research
Mission: Apply state-of-the-art research in AI, Machine Learning and Operations Research into Enterprise Products to automate workflows and enhance productivity.
Machine Learning
Classification
Regression
Personalization
Search Ranking
Recommendation
Anomaly Detection
Reinforcement Learning
Computer Vision
Object Detection
Person Identification
Video Search
Video Summarization
OCR
Natural Language Understanding
Text Classification
Entity Recognition
Text Summarization
Question Answering
Speech to Text
Natural Language Generation
Operations Research
Optimization
Scheduling
Network Planning
Product Mix
Inventory management
Resource Allocation
Enterprise Products
We are hiring in SF Bay Area, NYC, and Austin!


1.00
CREDIT
Background and Challenges
1/ Facebook extends credit to qualified advertisers to pay for ads.
Larger advertisers tend to use Extended Credit (Monthly Invoicing) vs. Self-Serve (Postpay).
2/ Extended Credit customer base growing rapidly
When customers approach their credit limit, we must decide whether to increase the limit to enable continued spend.
We need to make good decisions instantly under conditions of explosive volume growth.
Enterprise Products

CREDIT
Solution

In order to automate credit decisions, we created a machine learning model to predict customer default.

Our credit model is integral to the systems allowing FB to automate 1000s of credit decisions per month.

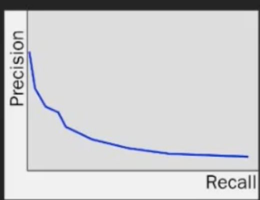


Enterprise Products

CREDIT
Learnings

1/ Most important metric takes into account dollar amounts

We use the bad debt capture curve in addition to standard ML metrics.



2/ Feature importance

Invoice payment history
Other on-platform behavior
Public financial data

3/ We can use our model to enable automation

More than half of customers have their credit decisions completely automated.

Most of the bad debt dollars are in the highest risk decile that we do not automate and review manually.

Enterprise Products

1.00

COLLECTIONS

Background and Challenges

1/ Not every customer pays their invoices on-time

Unlike automatic, pull payment methods (like credit cards), invoices are typically settled with a customer-initiated wire payment.

2/ Impractical & uneconomic to engage every customer as soon as they have overdues

To accommodate this, historically, we applied expert judgment and heuristics to determine which customers to prioritize engaging with to secure payment. For example, we would focus on customers with the most overdue invoices or longest overdue.

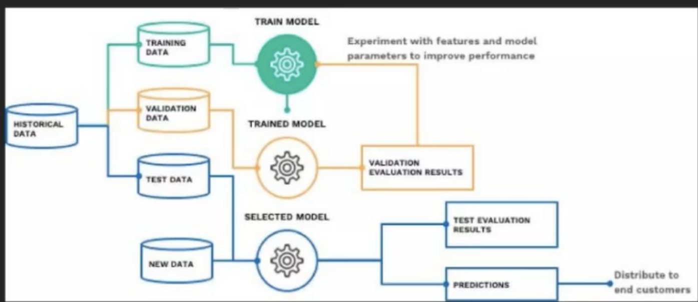
Enterprise Products

1.00

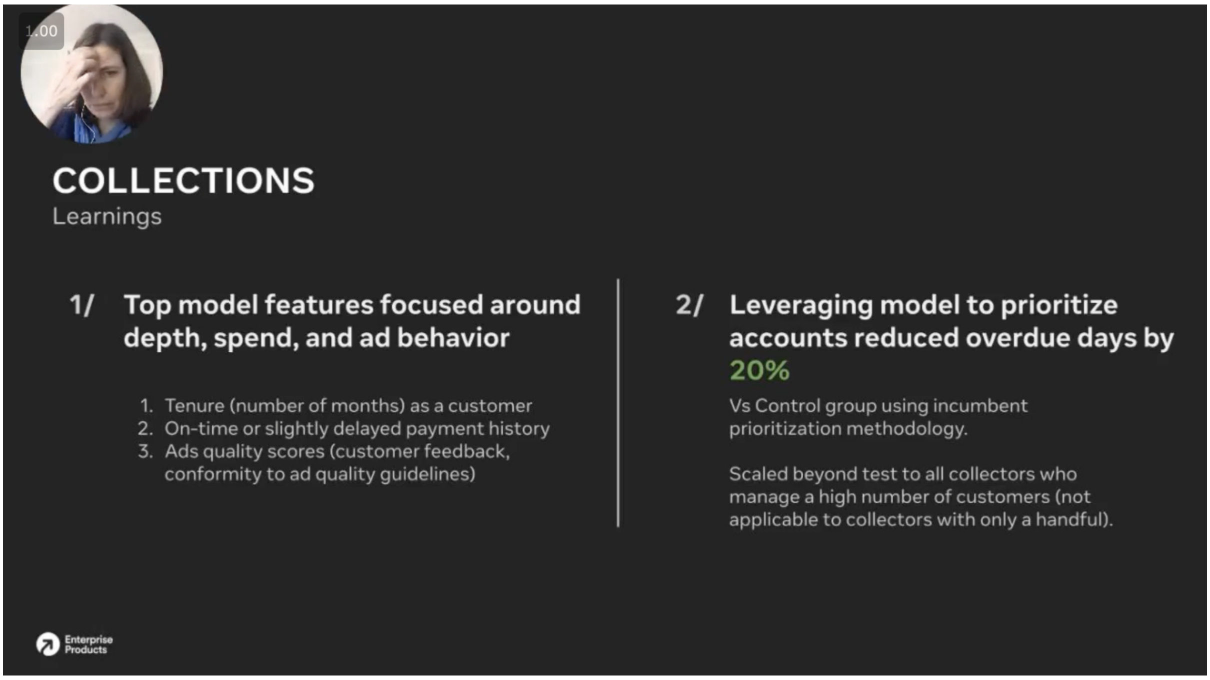
COLLECTIONS

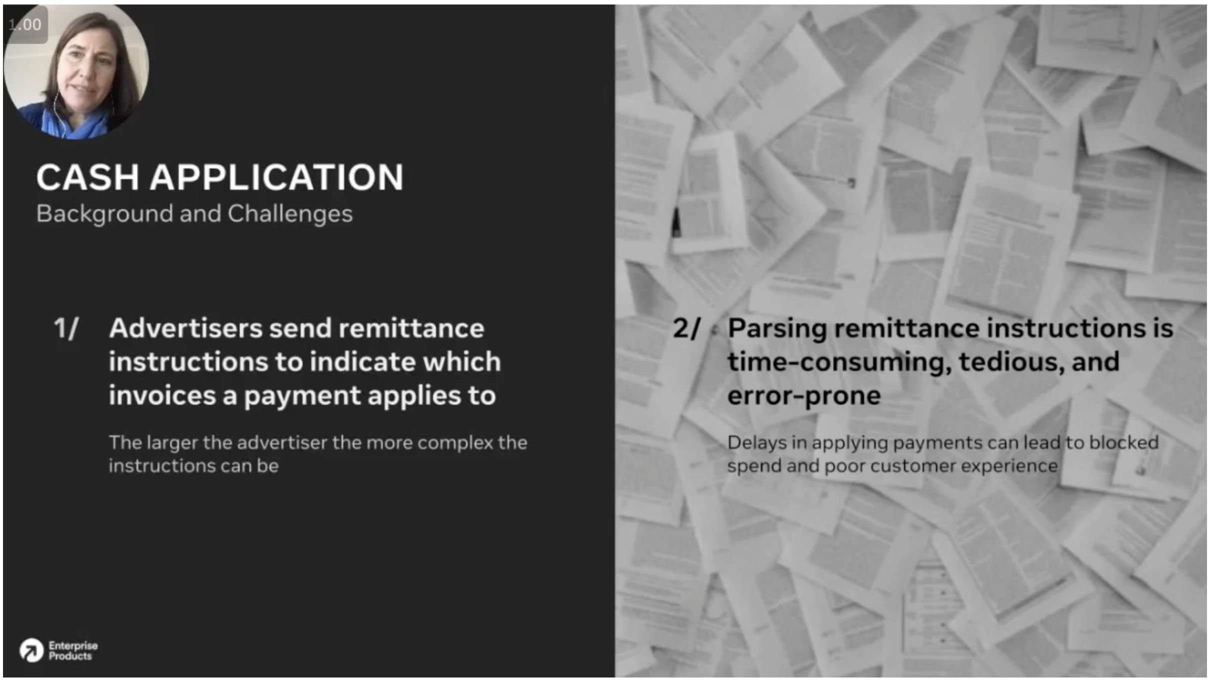
Solution

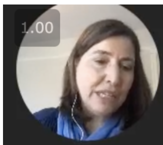
In order to identify customers with the greatest risk of non-payment, we created a machine learning model to predict invoices being unpaid / written off.



Enterprise Products


COLLECTIONS
Learnings
1/ Top model features focused around depth, spend, and ad behavior
1. Tenure (number of months) as a customer
2. On-time or slightly delayed payment history
3. Ads quality scores (customer feedback, conformity to ad quality guidelines)
2/ Leveraging model to prioritize accounts reduced overdue days by 20%
Vs Control group using incumbent prioritization methodology.
Scaled beyond test to all collectors who manage a high number of customers (not applicable to collectors with only a handful).
Enterprise Products

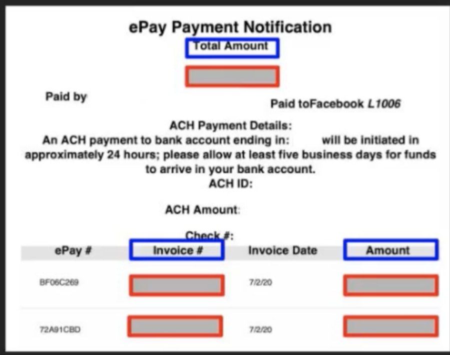

CASH APPLICATION
Background and Challenges
1/ Advertisers send remittance instructions to indicate which invoices a payment applies to
The larger the advertiser the more complex the instructions can be
2/ Parsing remittance instructions is time-consuming, tedious, and error-prone
Delays in applying payments can lead to blocked spend and poor customer experience
Enterprise Products



CASH APPLICATION

Solution

- Automatically reads remittances and applies cash accordingly
- Combines Optical Character Recognition (OCR) for info extraction with template-based document understanding
- For each template, we create a configuration describing fields of interest defined by name, value and spatial positioning



Enterprise Products

CASH APPLICATION

Learnings/Future Work

- Pareto effect in template coverage
- Significantly reduces average cycle time to apply payments
- Templates can change over time
- Ongoing work
  - UI for template creation
  - General ML solution

Enterprise Products

Lucas Peinado Bruscato

Do you usually use black box models to make the credit assessment or go with more tradicional ones?

Delete

Lucas Peinado Bruscato

David, in case you can talk about that, what type of models do you usually use?

Delete

Open 1 Answered 0

Lucas Peinado Bruscato

Do you model the customer or the transaction to decide if you are granting the credit?

Delete