

# How does it work?

## Facebook News Feed

CSE 705

Baasit Sharief (baasitsh@buffalo.edu)

June 6<sup>th</sup>, 2022

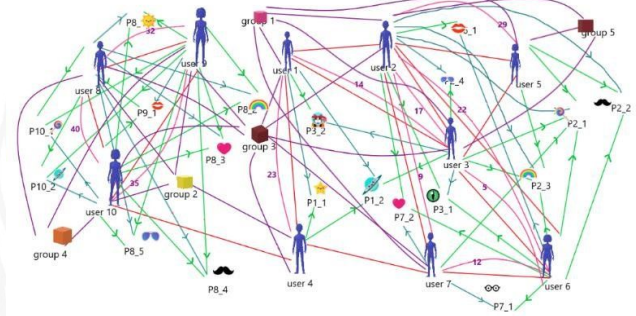


# Background



# Social Network and News Feed

- As of December 2017 more than two billion people users on Facebook
- News Feed is tailored to every individual to deliver content
- According to Facebook, News Feed shows you the stories that are most relevant to you.
- Here comes a question: how does Facebook define *“relevance”*?



## Brief History of News Feed

- First released in 2006, using popularity algorithms
- In 2007, '**like**' button was introduced on statuses/posts and comments
- Used the original chronological default ordering of posts depending on page liked and friends
- In 2009, facebook updated it to EdgeRank algorithm
- Edge weight calculation algorithm was updated until 2015
- EdgeRank was replaced with more accurate ML algorithms



# EdgeRank

- EdgeRank is like a credit rating: it's invisible, it's important, it's unique to each user, and no one other than Facebook knows exactly how it works.
- Three important aspects: **Affinity Score**, **Edge Weight** and **Time Decay**
- Affinity Score means how "connected" a particular user is to the Edge.
- Affinity score depends on user actions,
  - the strength of the action
  - how close the person who took the action was to you
  - how long ago they took the action.
- Every post on facebook is given some Edge weight i.e. their importance
  - Photos and videos have a higher weight than links
- As a story gets older, it loses points because it's "old news"
- EdgeRank is a running score—not a one-time score.
  - When a user logs into Facebook, their newsfeed is populated with edges that have the highest score at that very moment in time.
  - Your status update will only hit the newsfeed if it has a higher score—at that moment in time—than the other possible newsfeed stories.

# Why Change EdgeRank?

- Money
- Yes, it was all for Money
- In this era of attention economy, Facebook wanted to exploit high technology
  - to hold people's attention on their platform
  - More people means more money from advertisers
- Attention economy focuses on how people's limited attention is allocated among content (Simon et al., 1971)
- Increase user engagement to increase revenue through advertisements



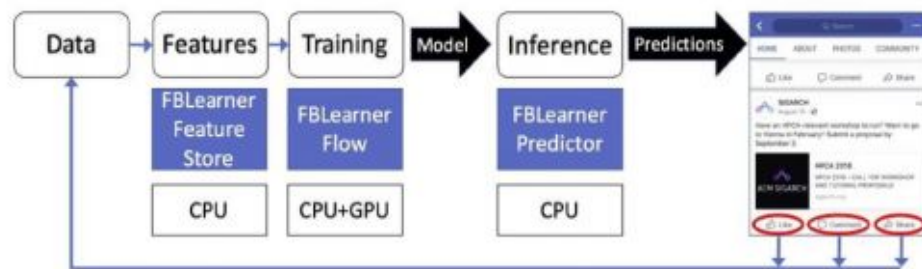


# Advent of ML and News Feed



# ML in Facebook

- Exploiting ML for better user engagement and interaction
- Algorithm involving has 4 parts, Inventory, Signals, Predictions, Scoring
- User engagement measured by various factors
  - View
  - Click
  - Like
  - Comment
  - Share
  - and 100000 more features
- Inputs are various content on Facebook, and the output is the probability of an engagement event
- Model generates a personalized set of relevant posts, images, and other content to display, as well as the best ordering of the chosen content

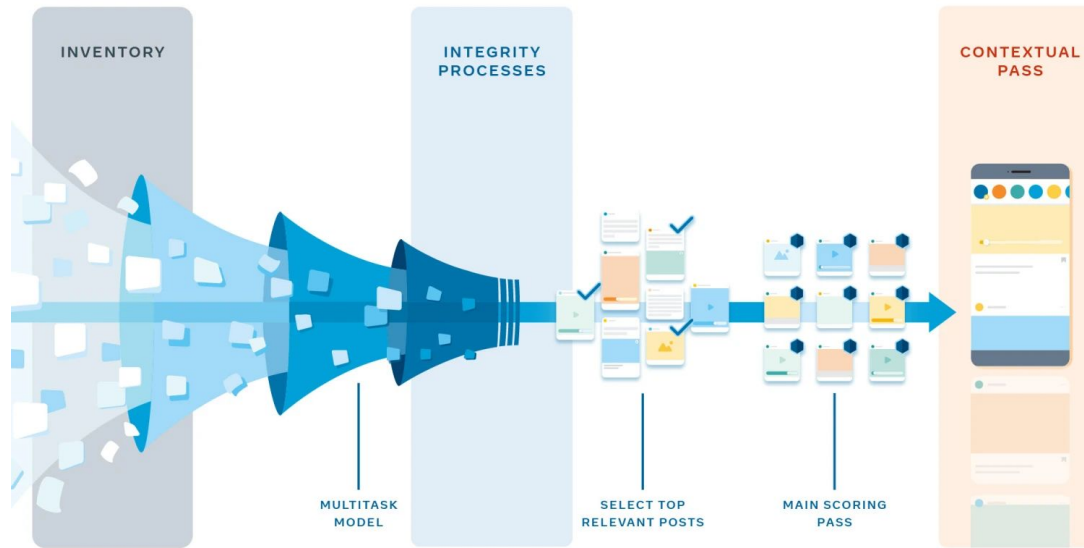


[2] [Deblackbox Facebook News Feed Algorithm as A System for Attention Manipulation](#)

[4] [Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective \(Hazelwood et al., 2018\)](#)

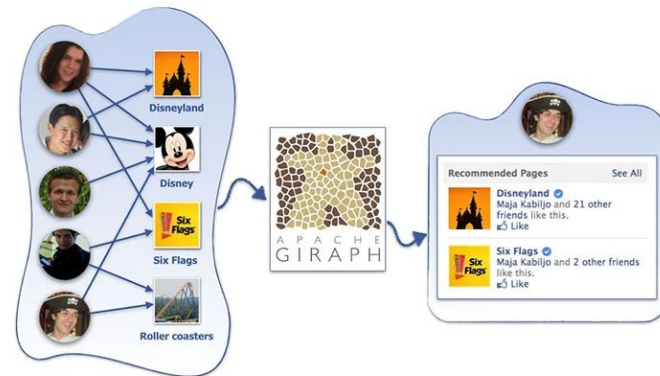


# Overview of the algorithm



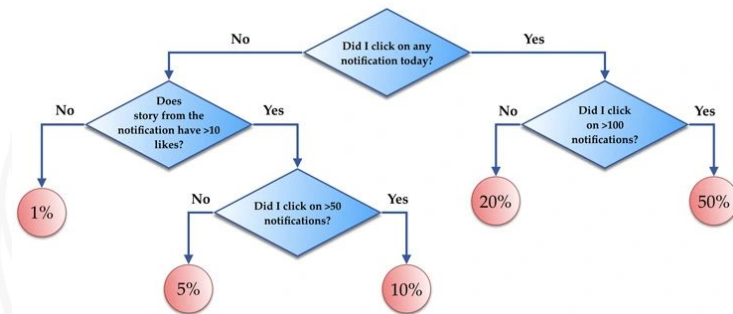
# Collaborative filtering

- Uses Apache Giraph to analyze the social graph formed by users and their connections
- Apache Giraph is an iterative graph processing system built for big data
- Recommendation system which gives suggestions based on the content people engage most with
- Three types - ***User-based collaborative filtering, Item-based collaborative filtering, and Model-based collaborative filtering***
  - User-based first finds neighbours who share similar interests and suggests posts preferred by their neighbours
  - Item-based calculates the similarity score between two posts based on all users' reaction to them and recommends based on user's preference
  - Model-based uses posts extracted from targeted users' prior reactions, to predict their future actions given a specific post
- Uses all 3 at unison to get best results



# Decision Trees for Engagement Prediction

- One of the oldest methods in Machine Learning
- Based on the idea that user's future action is consistent with their past actions
- Flowchart-like structure
  - each decision node represents a class label
  - each branch represents the outcome that leads to class labels
- Paths from root to leaves represent classification rules
- Leaves represent the decisions or final outcomes

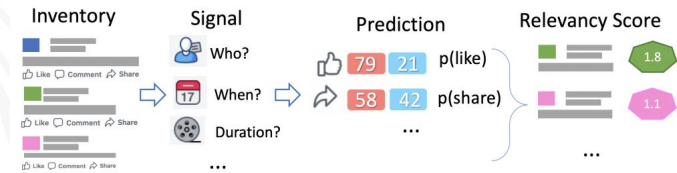


The probability of clicking on a notification

# Ranking/Scoring

- Calculates ranking score based on two aspects
  - Probability, engagement probability calculated by the DT model
  - Value, represents the weight given to an event
- Linear operation
  - Summation of probability multiplied with corresponding weights
- In 2016, Facebook stated core values
  - Post from friends and family first
  - Added “See First” option to users to give them more leverage
- Each user might have different value matrix unknown to the user based on which scoring is determined,
  - For example, for people who like to express themselves more through liking than commenting, the value of like is higher

Event	Probability	Value
Click	11%	1
Like	2.2%	5
Comment	0.41%	20
Share	0.054%	40
Friend	0.0062%	50
Hide	0.099%	-100
Total		0.2277

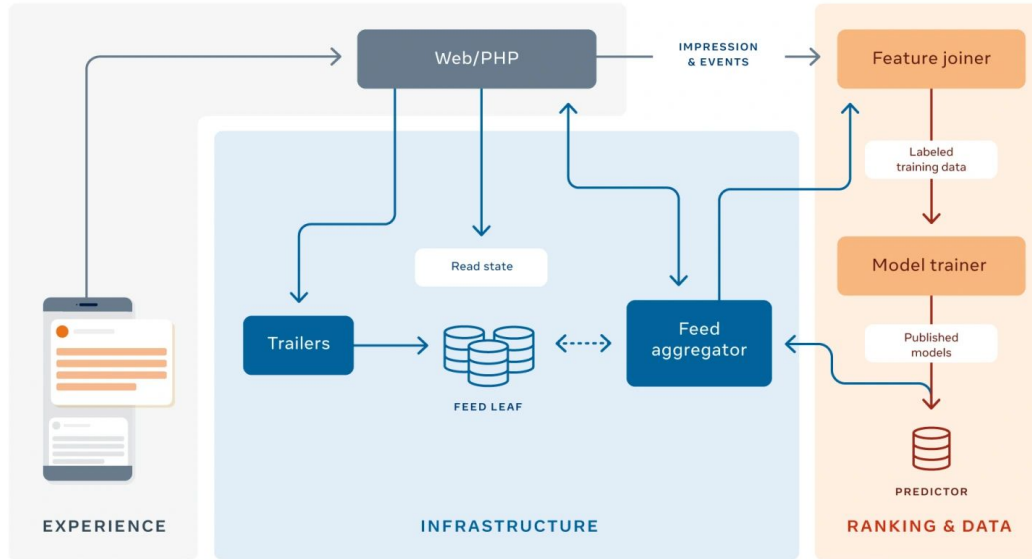


[2] [Deblackbox Facebook News Feed Algorithm as A System for Attention Manipulation](#)

[5] [News Feed ranking, powered by machine learning - Engineering at Meta \(fb.com\)](#)

[8] [Feature Selection for Facebook Feed Ranking System via a Group-Sparsity-Regularized Training Algorithm](#)

# Inference Infrastructure



When someone opens up Facebook, regardless of the front-end interface (e.g., iPhone, Android phone, web browser), the interface will send a request to a Web/PHP (front-end) layer, which then queries the feed aggregator (back-end layer). After accepting a request from the front end, the feed aggregator fetches actions and objects, along with an object summary, from the feed leaf databases so that it can process, aggregate, rank, and return the resulting list of ranked FeedStories to the front end for rendering.

## References

1. [A Study on Graph Theory Properties of Online Social Networks \(Lanel et al., 2020\)](#)
2. [Deblackbox Facebook News Feed Algorithm as A System for Attention Manipulation](#)
3. [What is EdgeRank?](#)
4. [Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective \(Hazelwood et al., 2018\)](#)
5. [News Feed ranking, powered by machine learning - Engineering at Meta](#)
6. [Recommending items to more than a billion people - Engineering at Meta](#)
7. [Evaluating boosted decision trees for billions of users](#)
8. [Feature Selection for Facebook Feed Ranking System via a Group-Sparsity-Regularized Training Algorithm](#)

# Thank You

