LIVE SESSION

GRAPH MACHINE LEARNING - INTRODUCTION AND LIVE DEMO

ON 23-JUL-2022 (SATURDAY)- ONLINE LIVE DELIVERY

TIME: 7 -8.30 PM IST (SATURDAY)
: 9.30 -11 AM EST (SATURDAY)
: 2.30-4PM LONDON (SATURDAY)

REGISTER FREE

HTTPS://FORMS.GLE/YH5FW1HHVDCCBE7J8

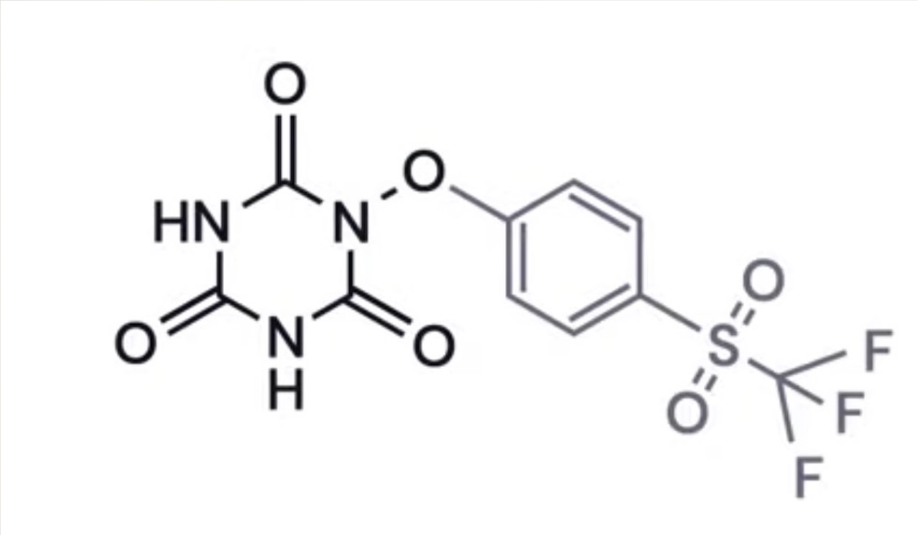
PRASANNA VENKATESH J

AI AND CLOUD EVANGELIST

GRAPH MACHINE LEARNING - AGENDA

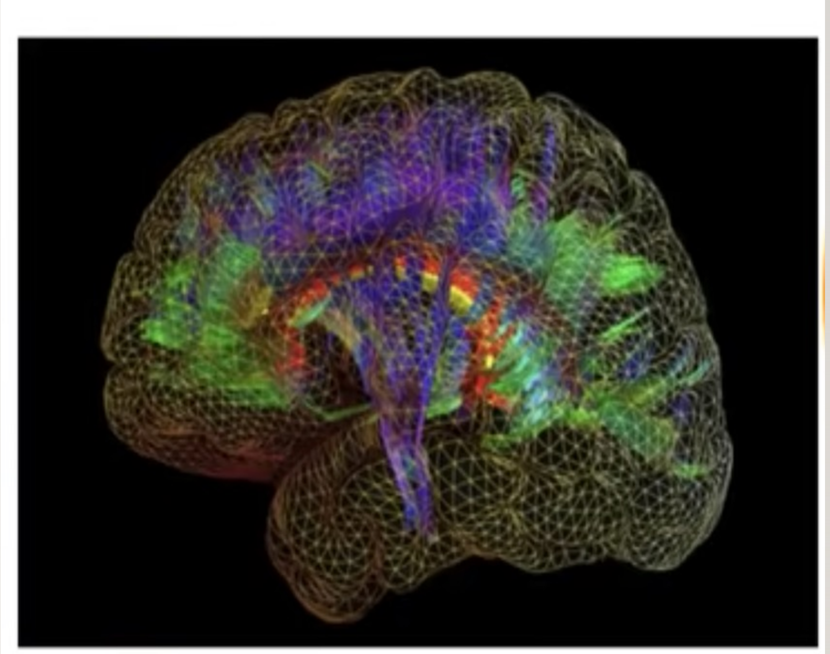
- Introduction to graph
- Industry framework and Graph storage options
- Introduction to Tiger Graph and features & DEMO
- Graph Neural Network Introduction and various frameworks
- Introduction to DGL (Deep Graph Library) and DEMO

INTRODUCTION TO GRAPH

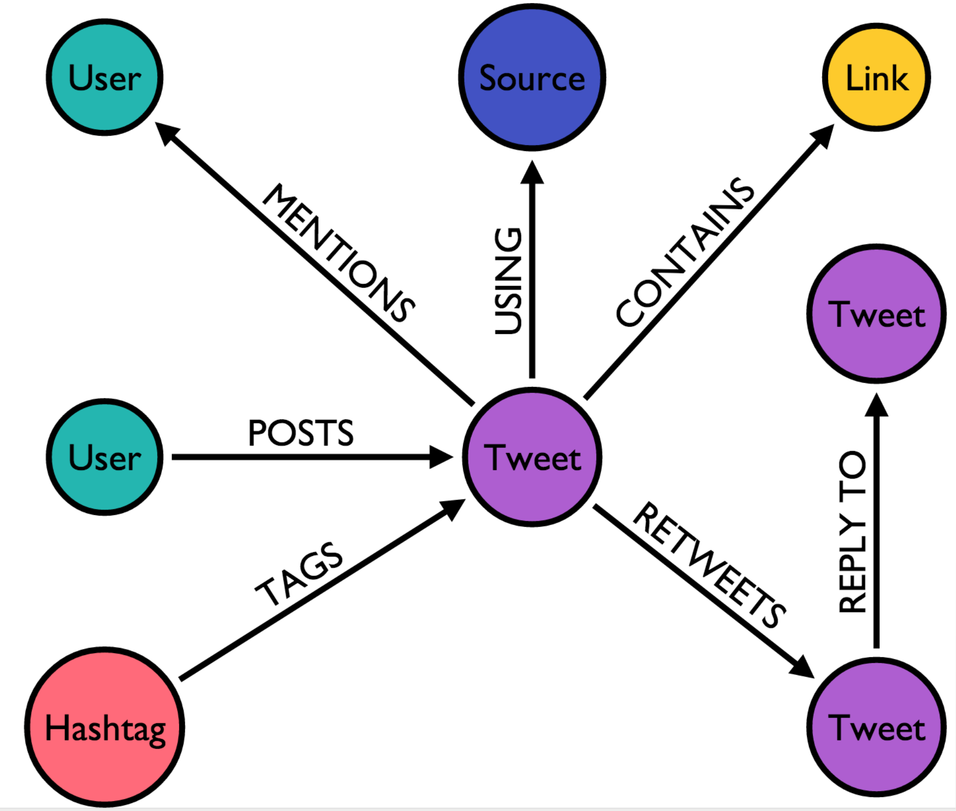


prasannavj@gmail.com

INTRODUCTION TO GRAPH



prasannavj@gmail.com



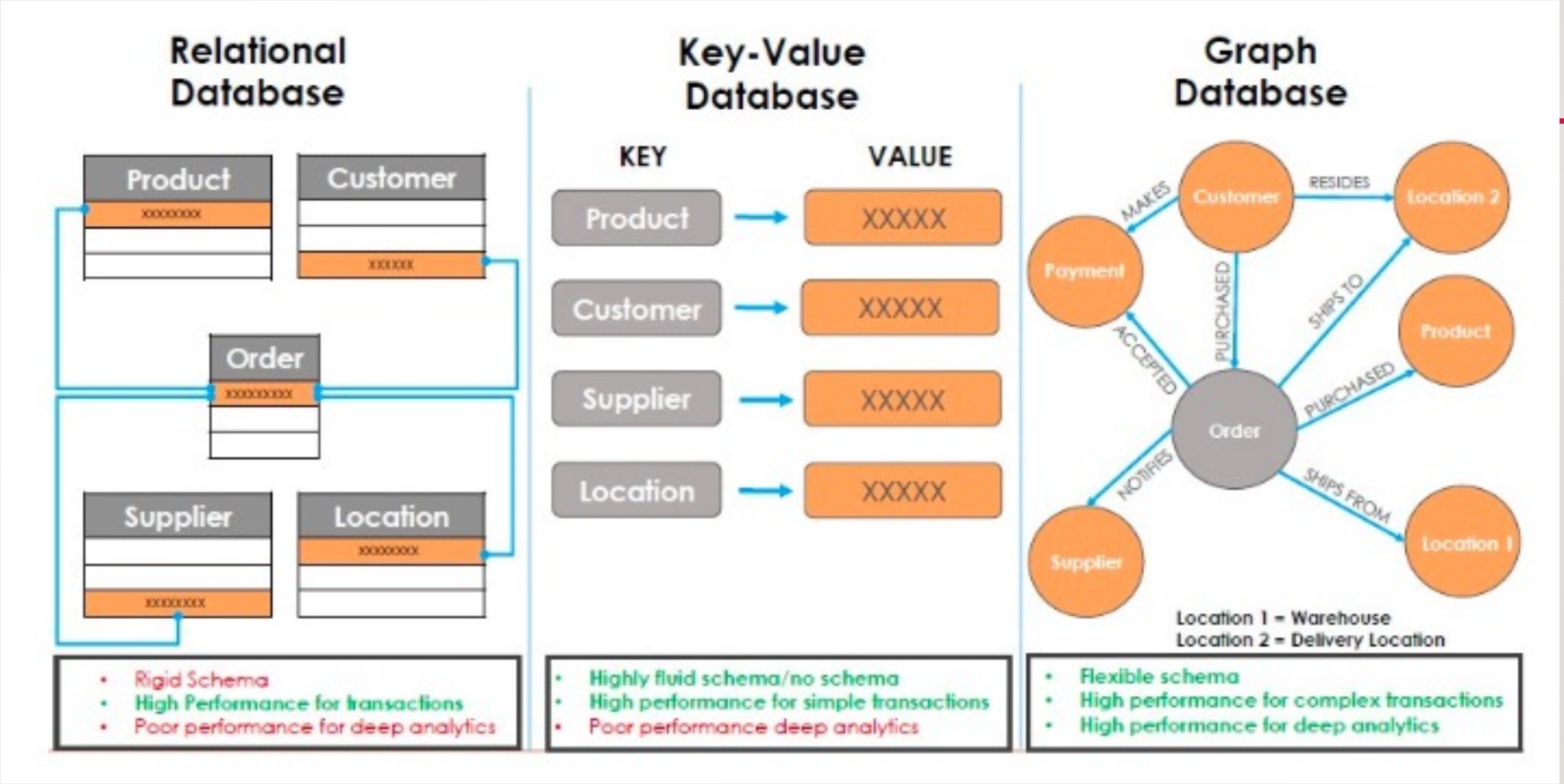
INTRODUCTION TO GRAPHS

- NODE
- EDGES

Typical Machine learning and deep learning frameworks focus on the Tabuler and related data.

Edges and the relationship has more meaningful information which are not processed with the current deep learning frameworks.

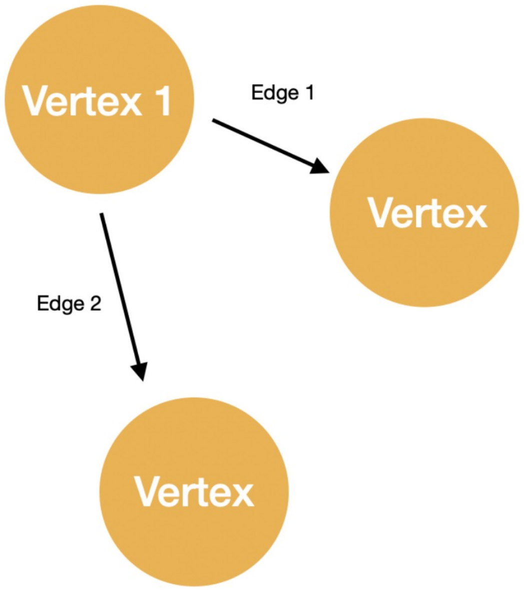
EVOLUTION OF DATABASES



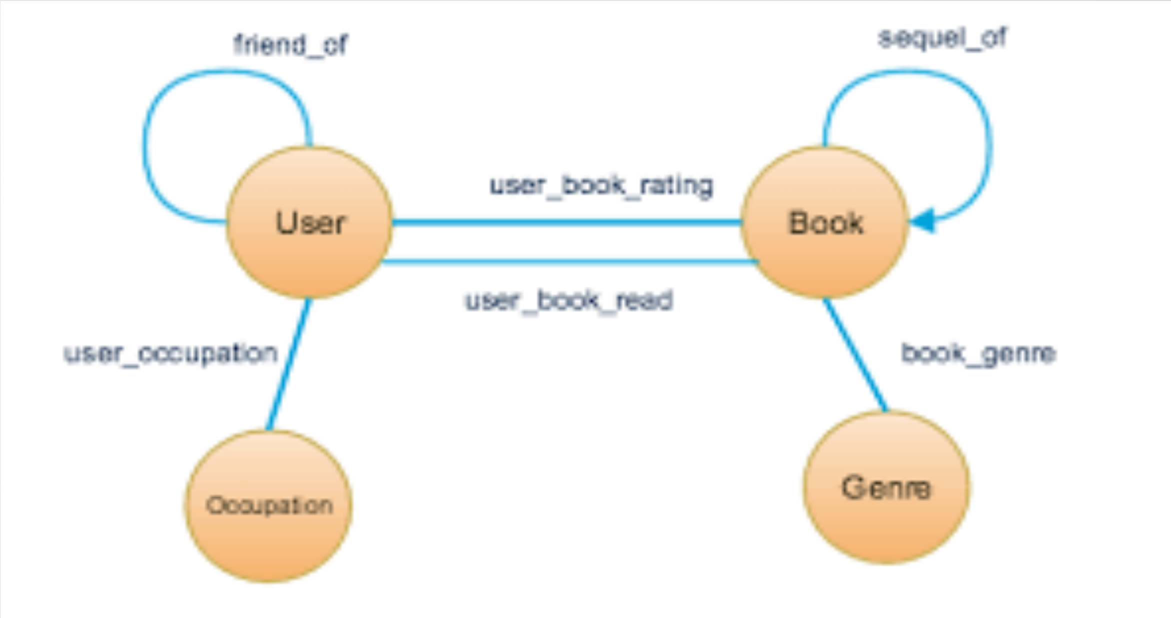
prasannavj@gmail.com

GRAPH DATA REPRESENTATION

Properties of Vertex 1
primary_id TYPE
Attribute1 TYPE
Attribute2 TYPE
Attribute3 TYPE

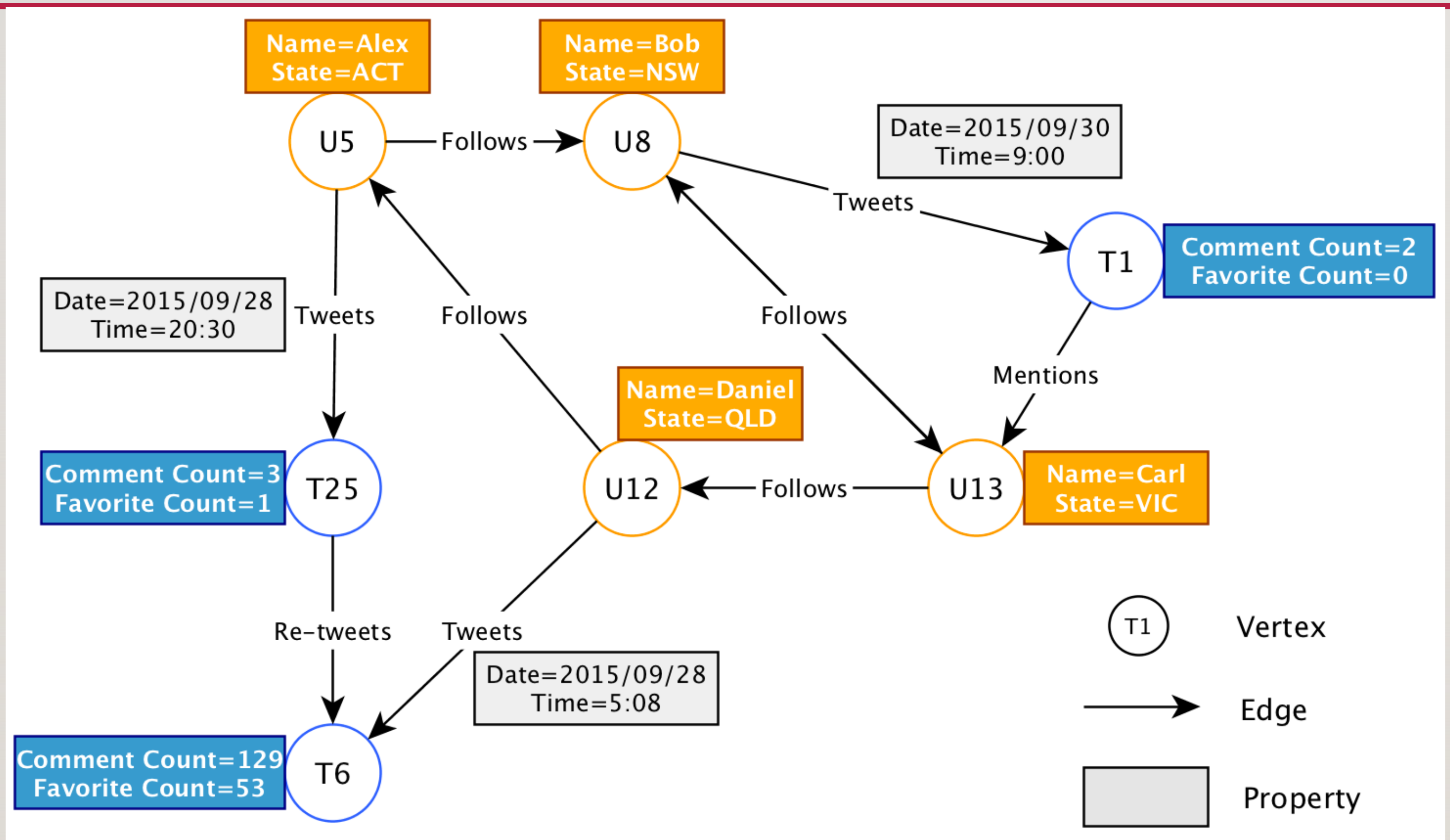


- Vertex to represent the Node
- Edge to connect different vertex
- Featurs to represent the Vertex and the Edge

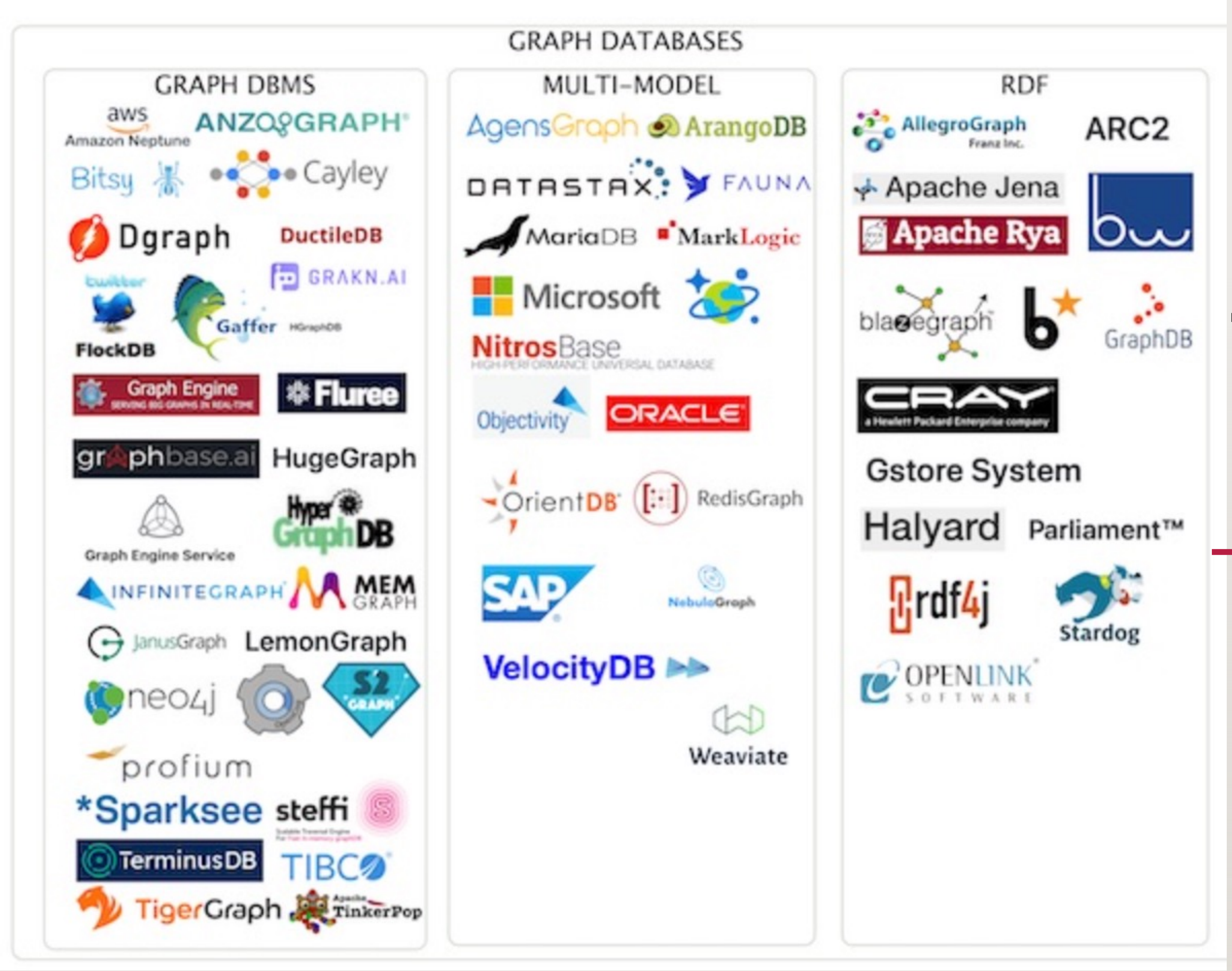


prasannavj@gmail.com

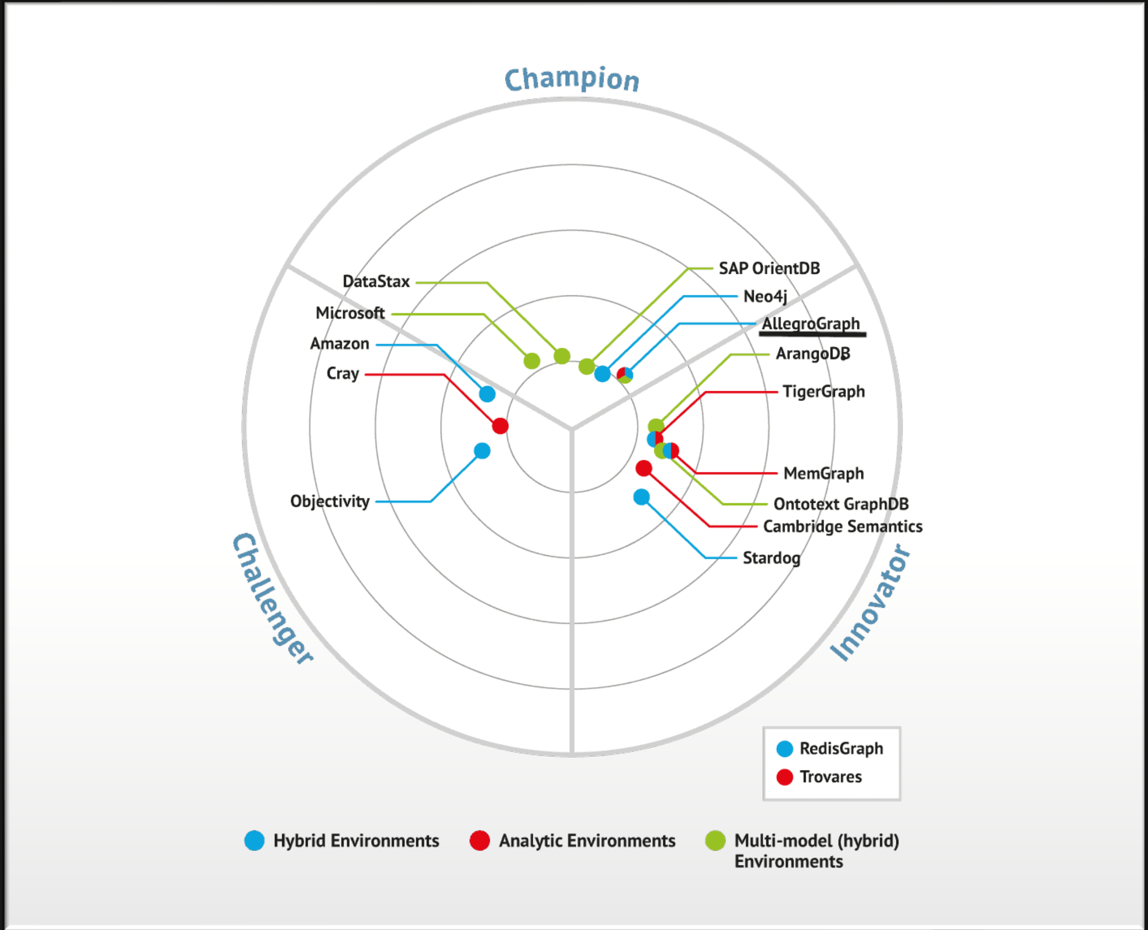
GRAPH DB INTRODUCTION



prasannavj@gmail.com



FRAMEWORKS TO REPRESENT GRAPH DATA



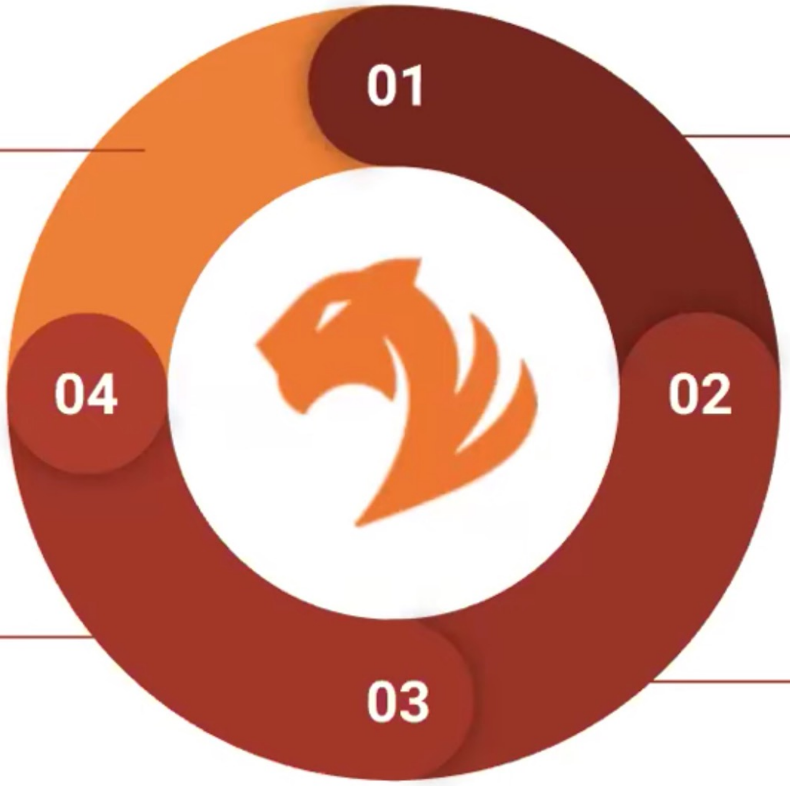
TIGERGRAPH A LEADING INNOVATOR

Source bloor research

The Right Models

ML Workbench

- Enabling supervised ML techniques with GNN on large data set
- Enabling vertex, link, graph prediction with a model learn directly on graph structure
- Fully Integrated with TigerGraph DB
- Plug and Play with customer's existing ML Pipeline and Infrastructure
- → More Graph analytical power for our customers



The Right Platform

TigerGraph DB

- MPP, Distributed Graph Database
- Scalability for massive data set
- Real time performance
- Enterprise grade security
- Graph Studio visual SDK

Graph Data Science Lib

The Right Algos

- 55+ Graph and Unsupervised Graph ML algorithms for graph and deep link analysis
- Graph feature generation for ML models.
- Performant, customizable, open-source, in-DB learning

GSQL

The Right Language

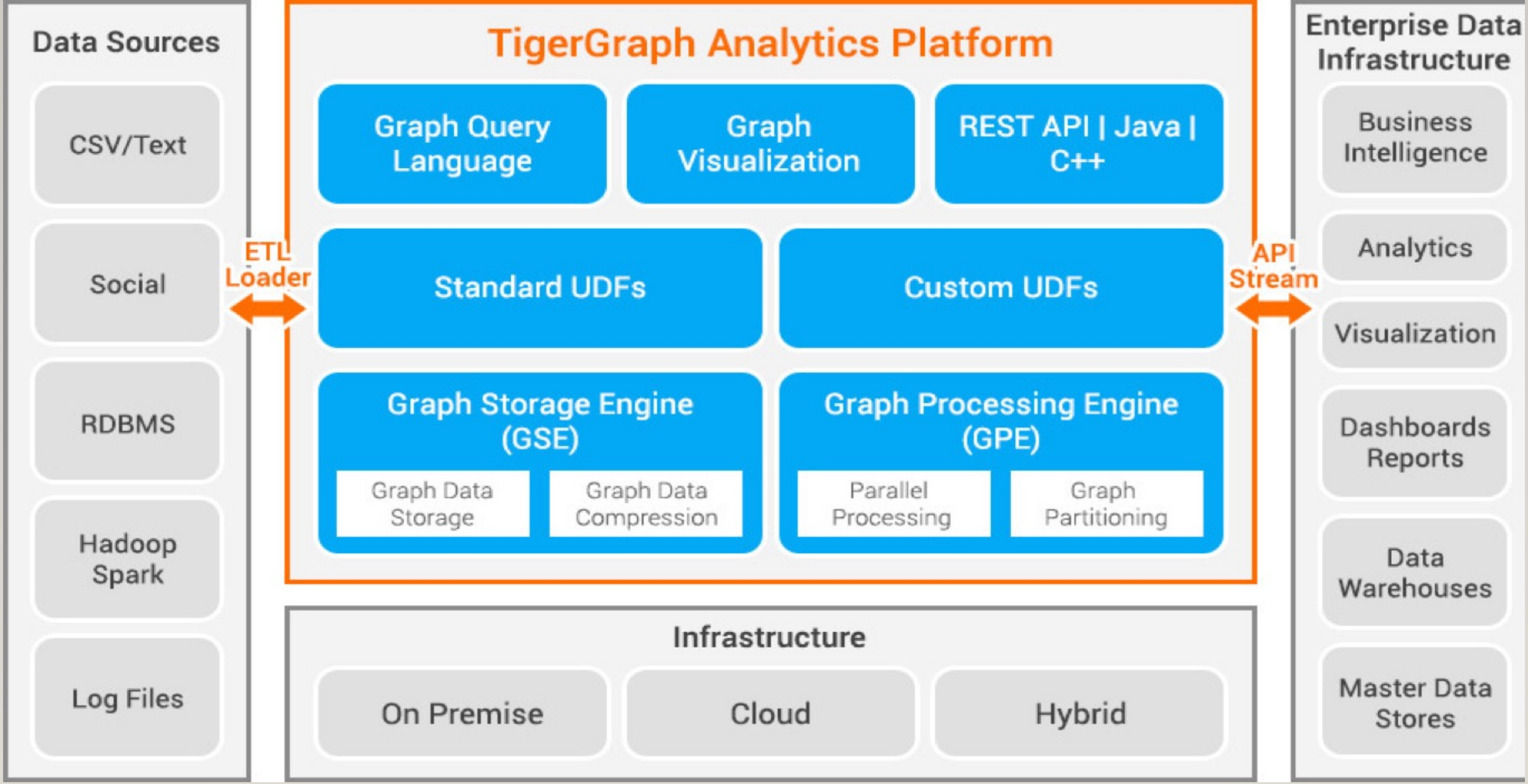
- Intuitive with a familiarity to SQL
- Turing complete graph query language
- Built in parallelism
- Built in accumulation

prasannavj@gmail.com

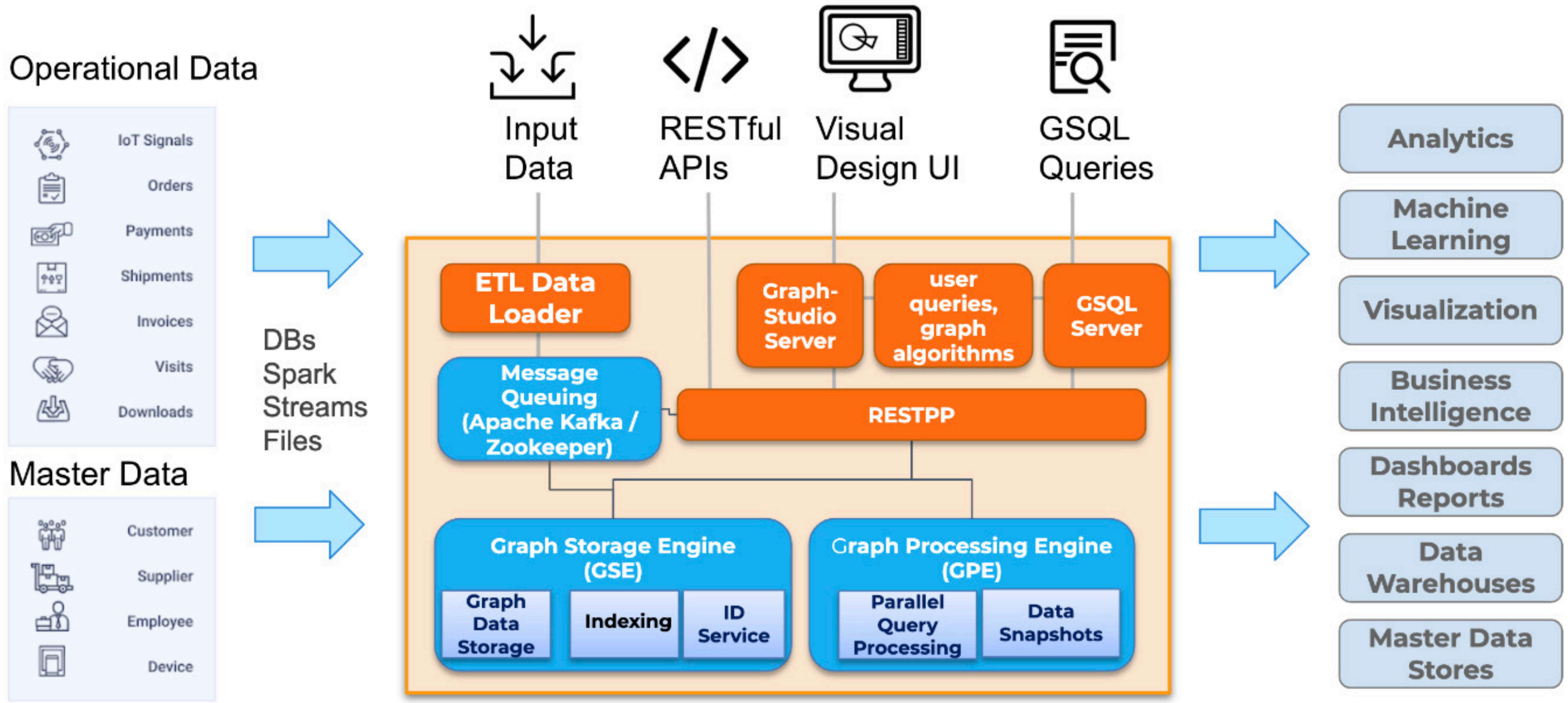
TIGER GRAPH ADVANTAGE

- A complete, distributed, parallel graph computing platform supporting web-scale data analytics in **real-time**.
- MPP – Massive parallel processing included.
- Support for multihop query which can work well beyond 2 hop.
- Cloud based access

prasannavj@gmail.com



prasannavj@gmail.com



prasannavj@gmail.com

TIGER GRAPH STUDIO - WORKFLOW

Design Schema → Map data to Graph → Load Data → Query Data → Visualize and analyze

prasannavj@gmail.com

TIGER GRAPH DEMO

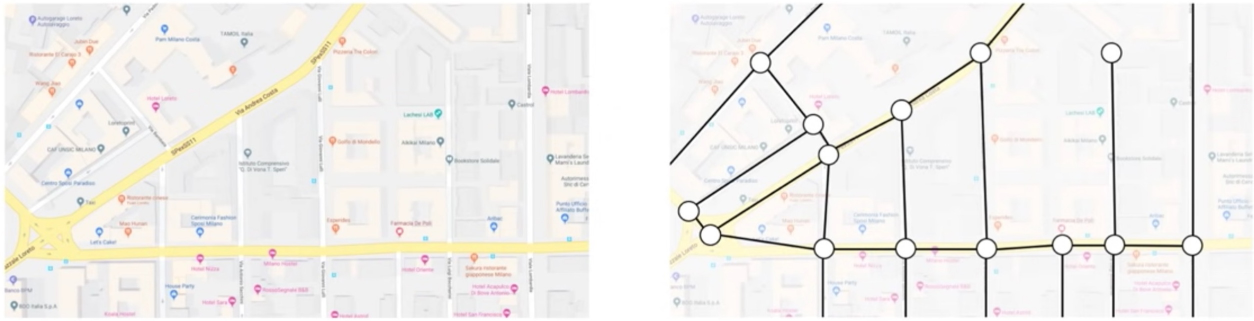
https://testdrive.tigergraph.com/app/home

prasannavj@gmail.com

GRAPH MACHINE LEARNING - INTRO
GOOGLE MAPS USING GRAPH FROM DEEPMIND

Traffic maps are graphs!

Transportation maps (e.g. the ones found on *Google Maps*)
naturally modeled as **graphs**.



Nodes could be **intersections**, and edges could be **roads**.
(Relevant **node features**: road *length, current speeds, historical speeds*)

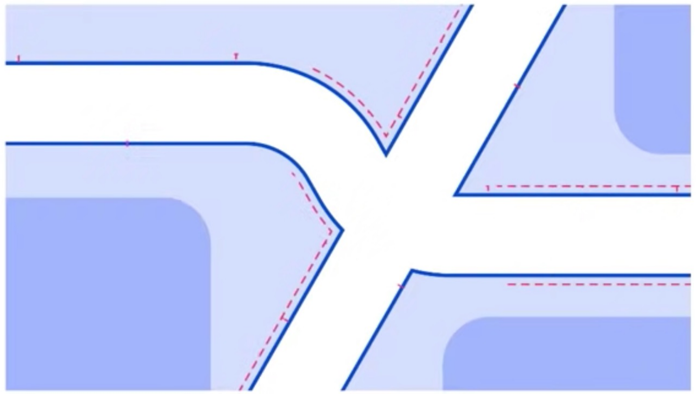
prasannavj@gmail.com

DEEPMIND ETA PREDICTIONS

DeepMind's ETA Prediction: GraphNets on Supersegments

Partition candidate route into **supersegments**, sampled proportionally to (est.) **traffic density**.

Run GNN on **supersegment** graph to estimate estimated time of arrival (ETA) (*graph regression*).

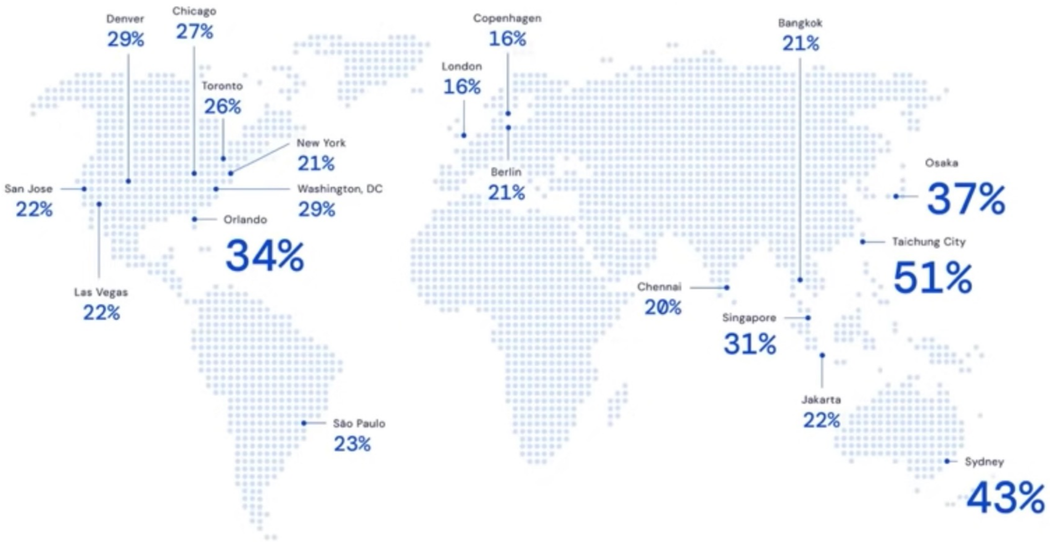


prasannavj@gmail.com

IMPROVEMENT IN ETA ON GOOGLE MAPS

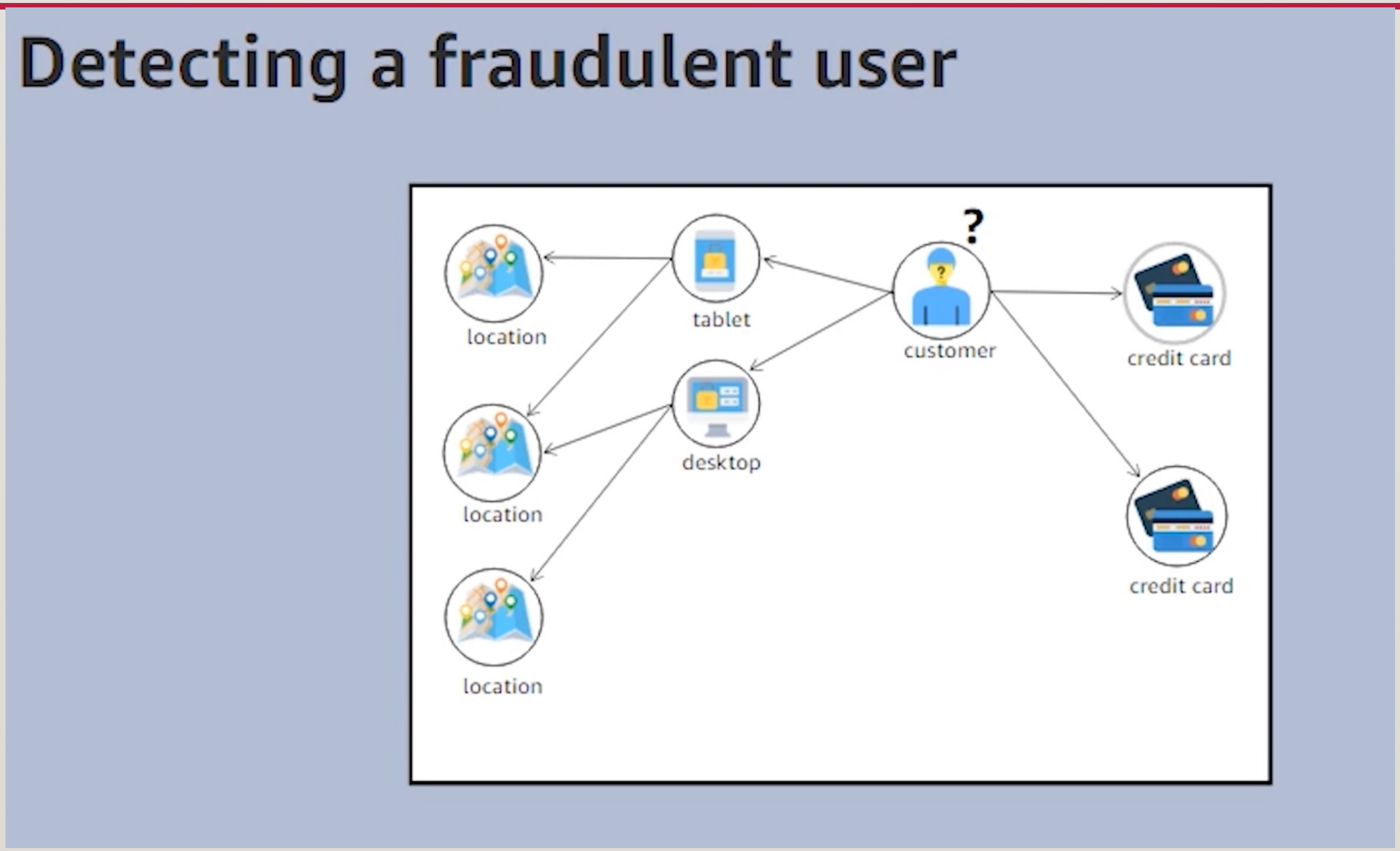
Returns

Already **deployed** in several major cities, significantly reducing negative ETA outcomes!



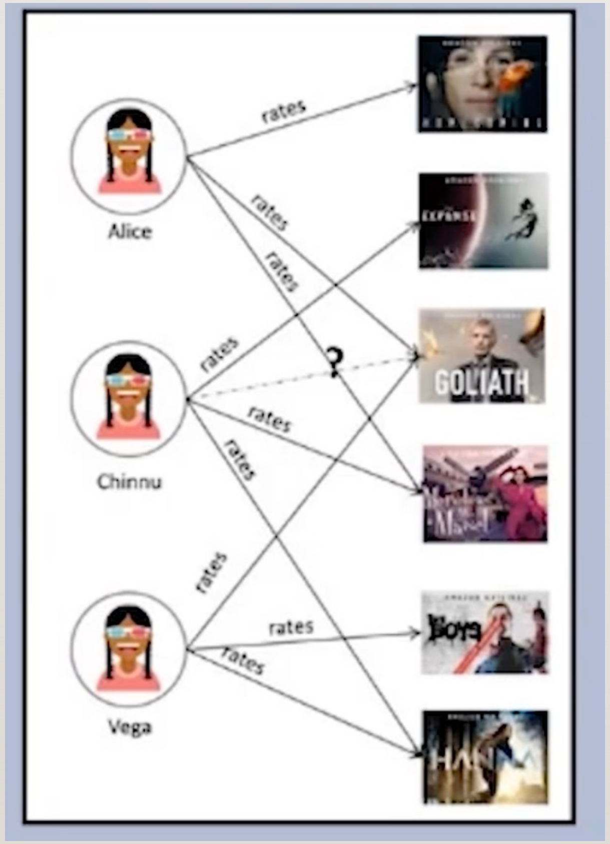
prasannavj@gmail.com

FRAUD DETECTION



prasannavj@gmail.com

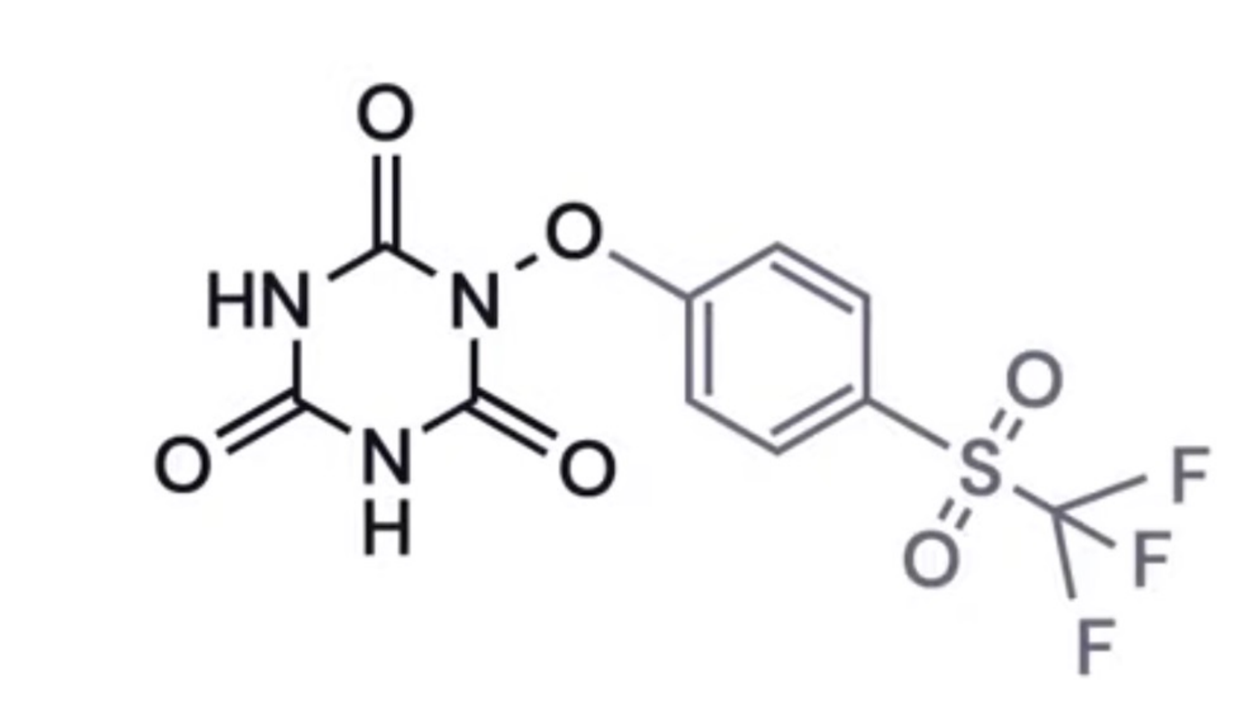
RECOMENDATIONS



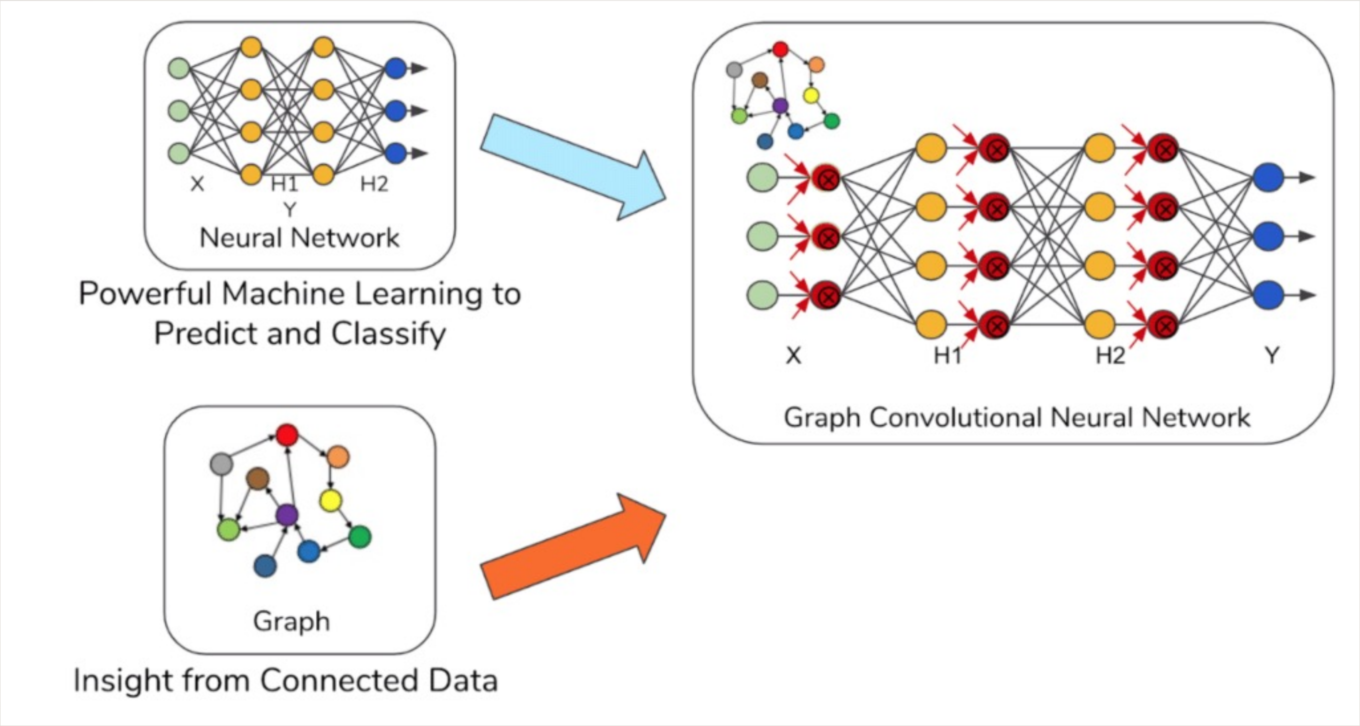
Should we recommend GOLIATH Movie to Chinnu??

prasannavj@gmail.com

DRUG DISCOVERY



TIGER GRAPH MACHINE LEARNING



prasannavj@gmail.com

Graph Data Science Library

Shortest Path

Find the shortest path or evaluate availability and quality of routes

Centrality

Determine importance of each vertex within a network

Community Detection

Discover clusters of high interconnection

Similarity

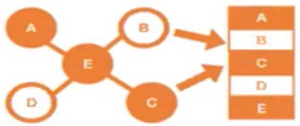
Compare features of vertices and/or their relationships

Classification

Predict the class of an entity, based on the evidence of previously classified entities

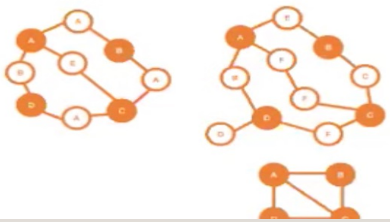
Topological Link Prediction

Predict the existence of a link between two entities



Graph Embeddings

Convert the neighborhood topology of each vertex into a fixed size vector of decimal values



Frequent Pattern Mining

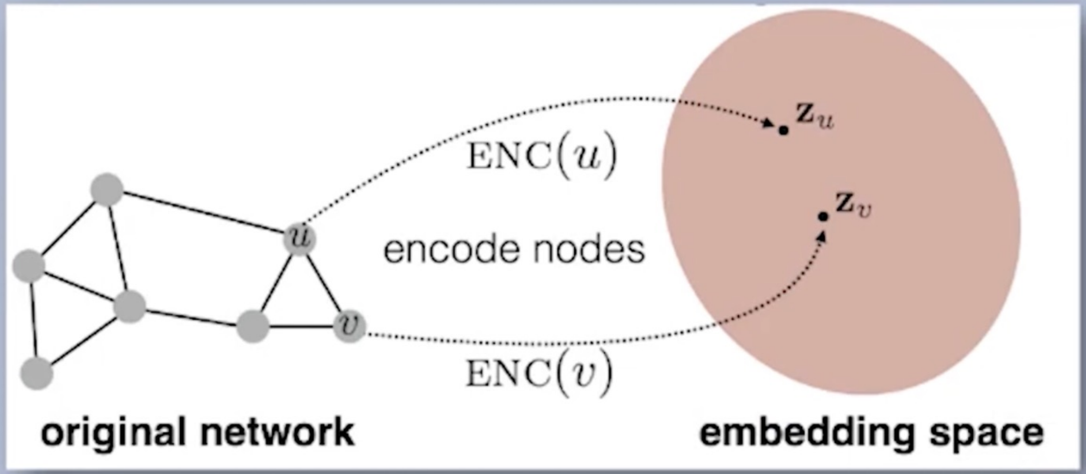
Find subgraph patterns that occur the most frequently

prasannavj@gmail.com

INTRODUCTION TO GRAPH NEURAL NETWORK

Graph learning and node embeddings

Embed nodes to a low-dimension space by capturing the essential task-specific information and use them to train off-the-self classifiers.

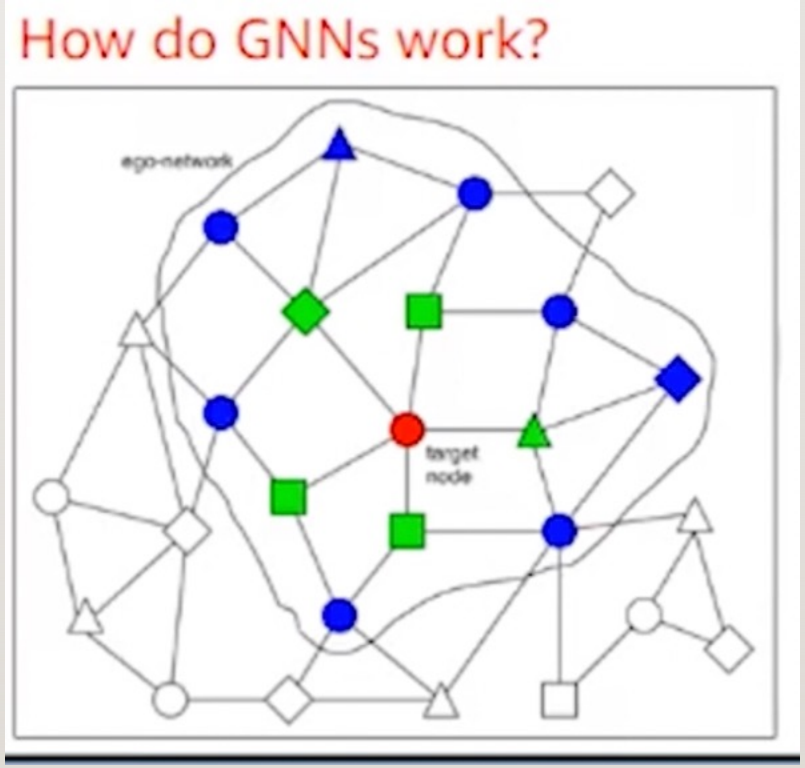


http://snap.stanford.edu/proj/embeddings-www/

prasannavj@gmail.com

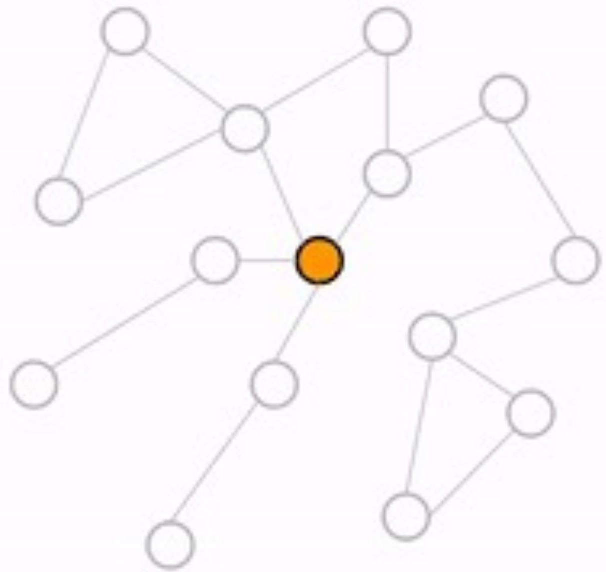
GRAPH NEURAL NETWORK (GNN)

A Family of (deep) neural networks that learn node, edge and graph embedding.



prasannavj@gmail.com

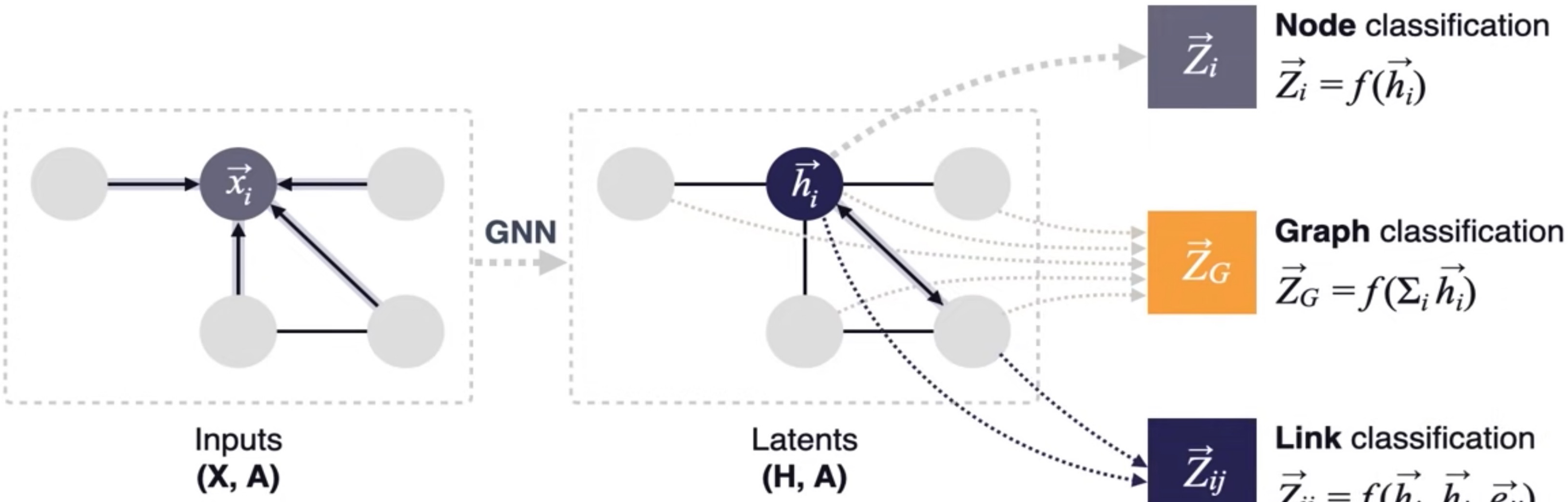
Input Layer



» Node 1 → One-hot vector [0,0,1,0,0]

prasannavj@gmail.com

What to do with GNN outputs?

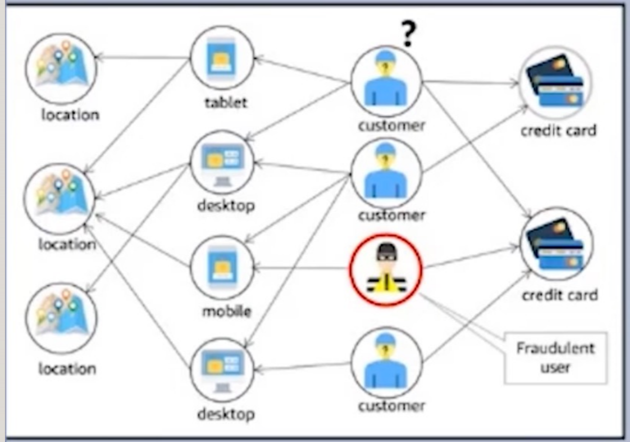


prasannavj@gmail.com

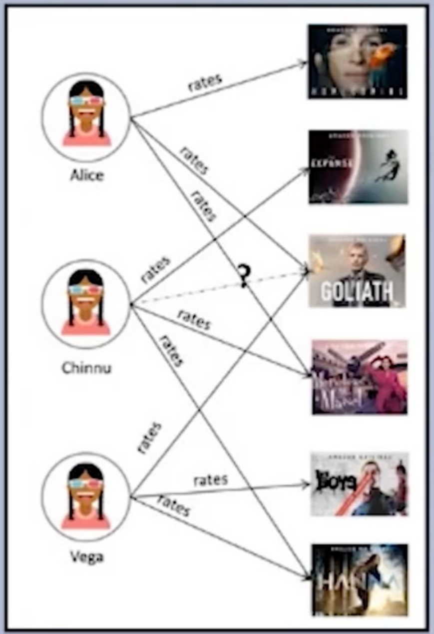
TYPES OF PREDICTIONS

Tasks in graph learning

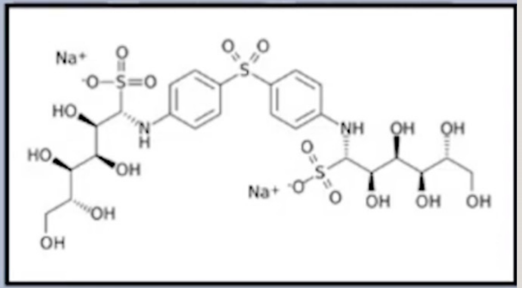
Node-level prediction



Edge-level prediction



Graph-level prediction

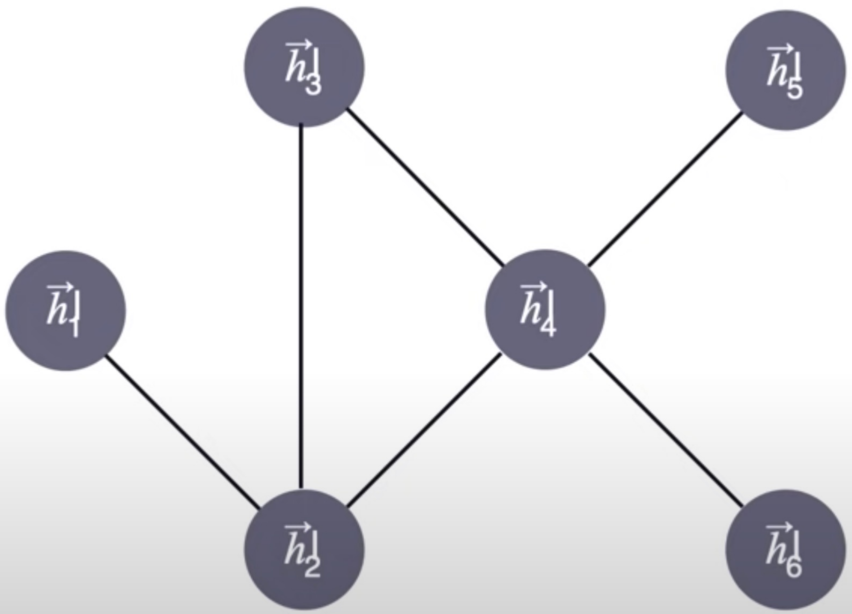


prasannavj@gmail.com

VARIENCE OF GRAPH NEURAL NETWORKS

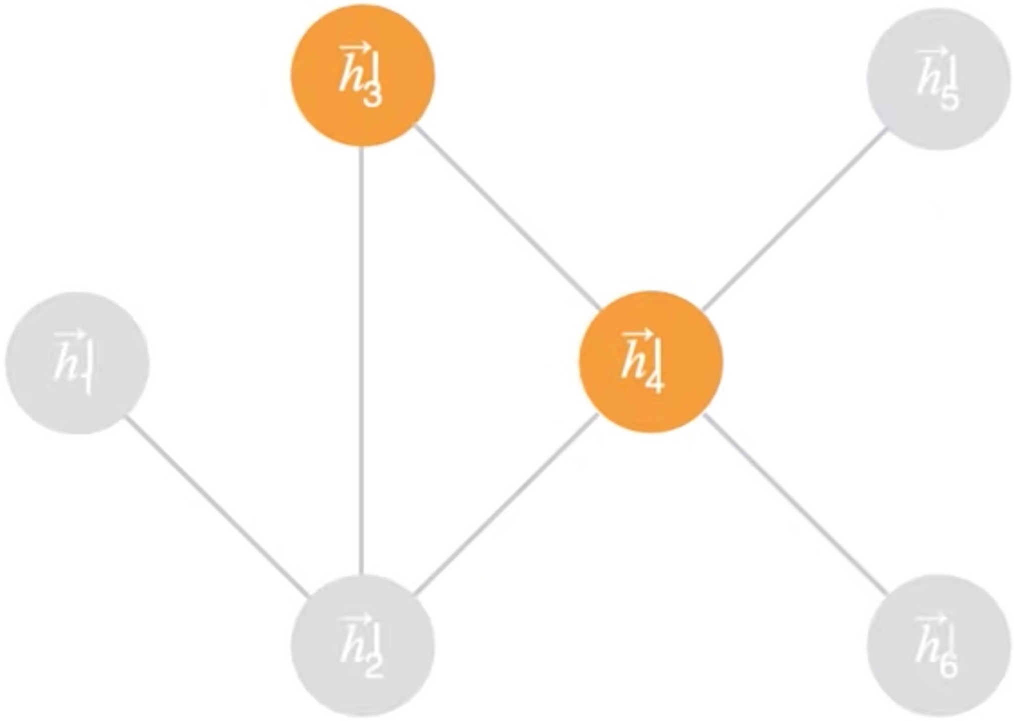
- MPNN – Message Passing Neural Network
- GraphSAGE (PinSAGE)

MPNN: Initial setup



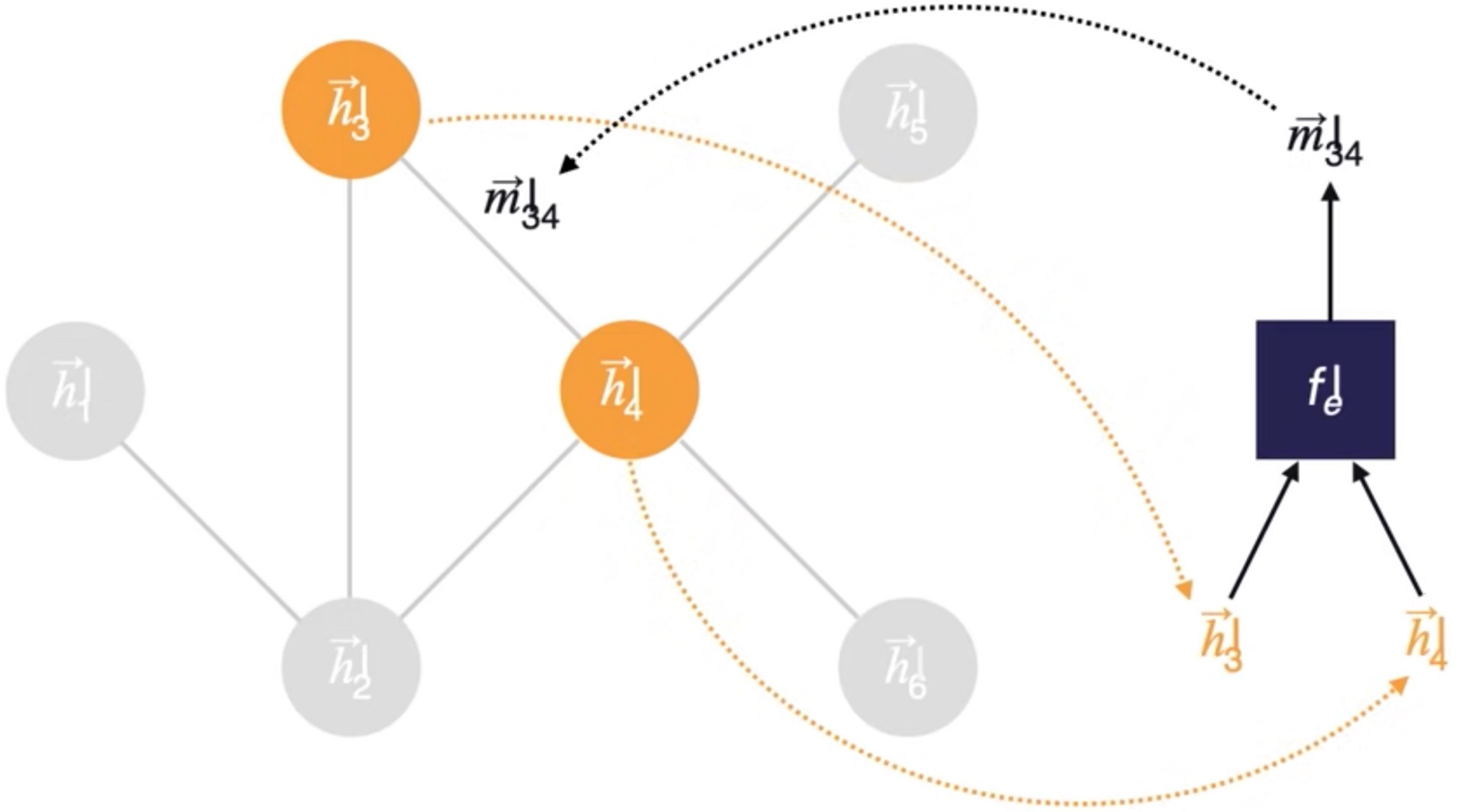
prasannavj@gmail.com

MPNN: Next-level features

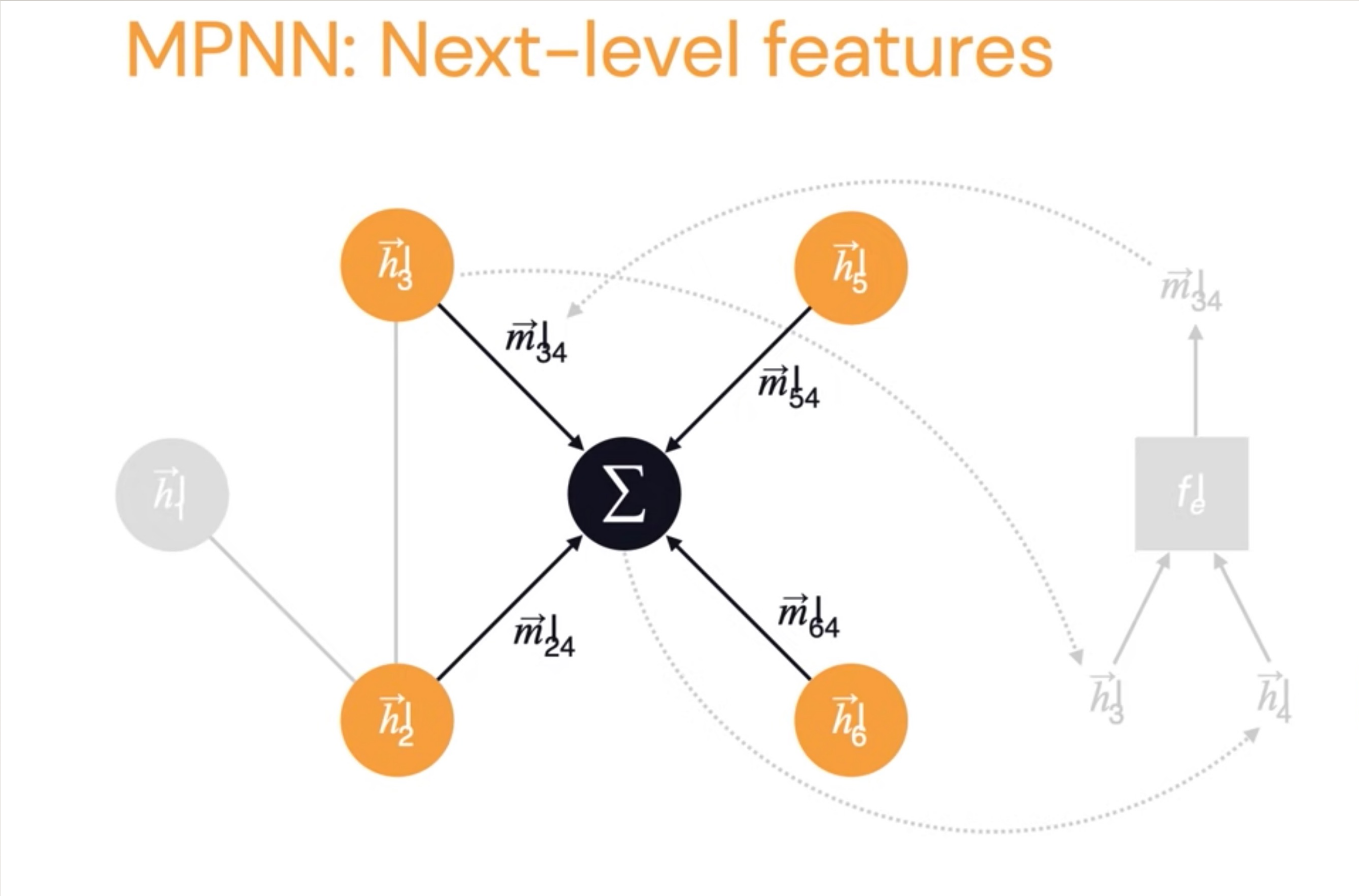


prasannavj@gmail.com

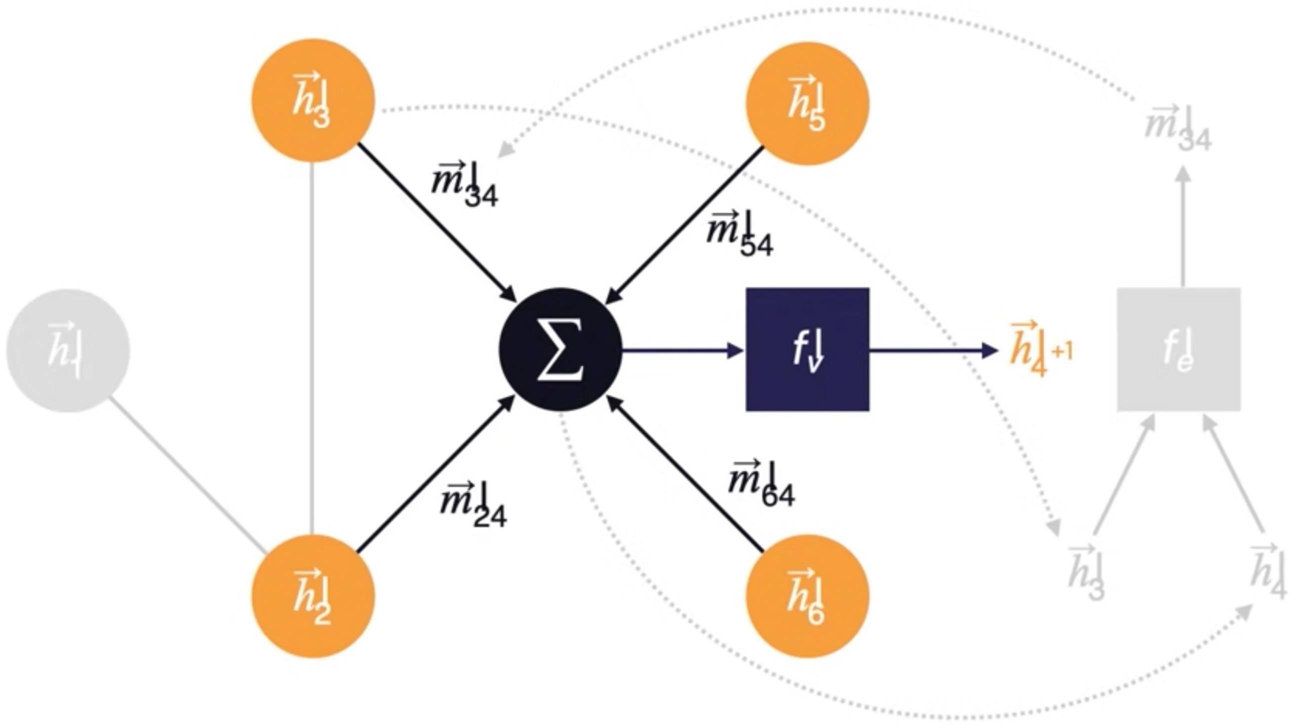
MPNN: Next-level features



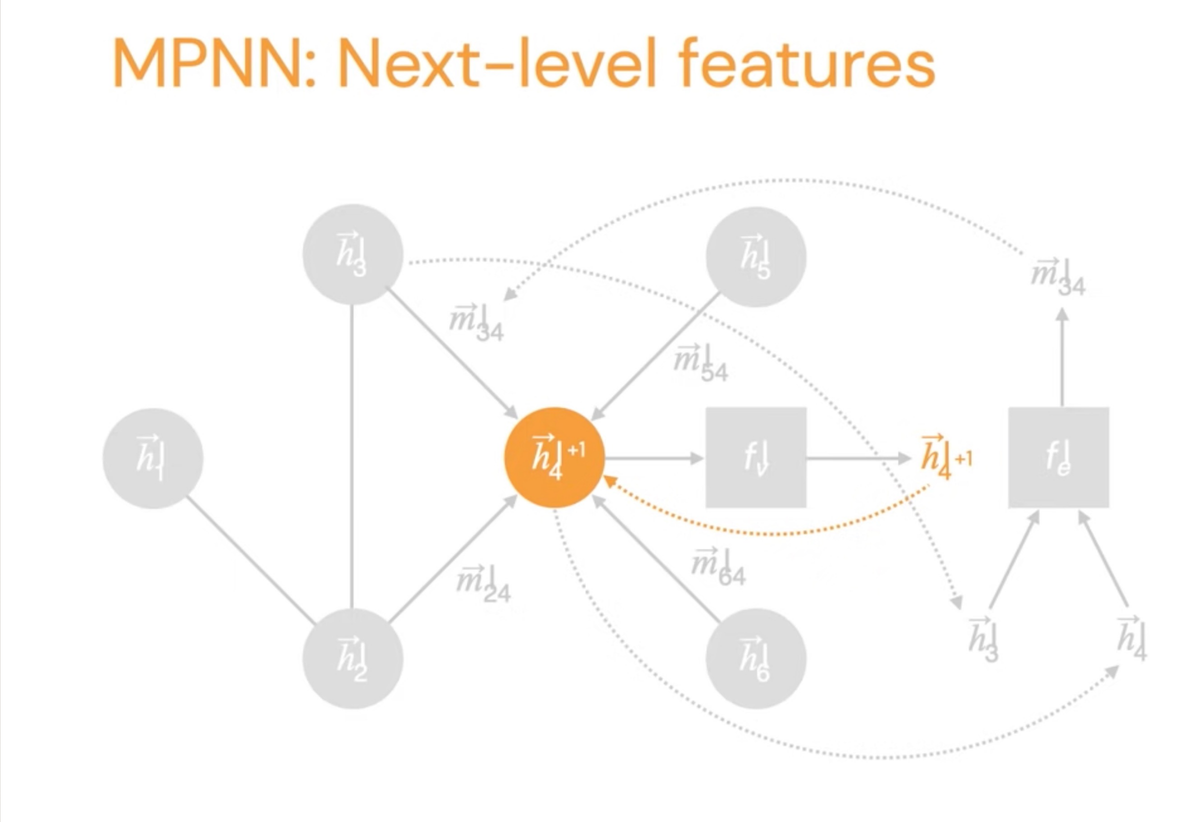
prasannavj@gmail.com


MPNN: Next-level features
$\vec{h}_3^l$
$\vec{h}_5^l$
$\vec{m}_{34}^l$
$\vec{m}_{54}^l$
$\vec{h}_1^l$
Σ
f_e^l
$\vec{m}_{24}^l$
$\vec{m}_{64}^l$
$\vec{h}_2^l$
$\vec{h}_6^l$
$\vec{h}_3^l$
$\vec{h}_4^l$

MPNN: Next-level features

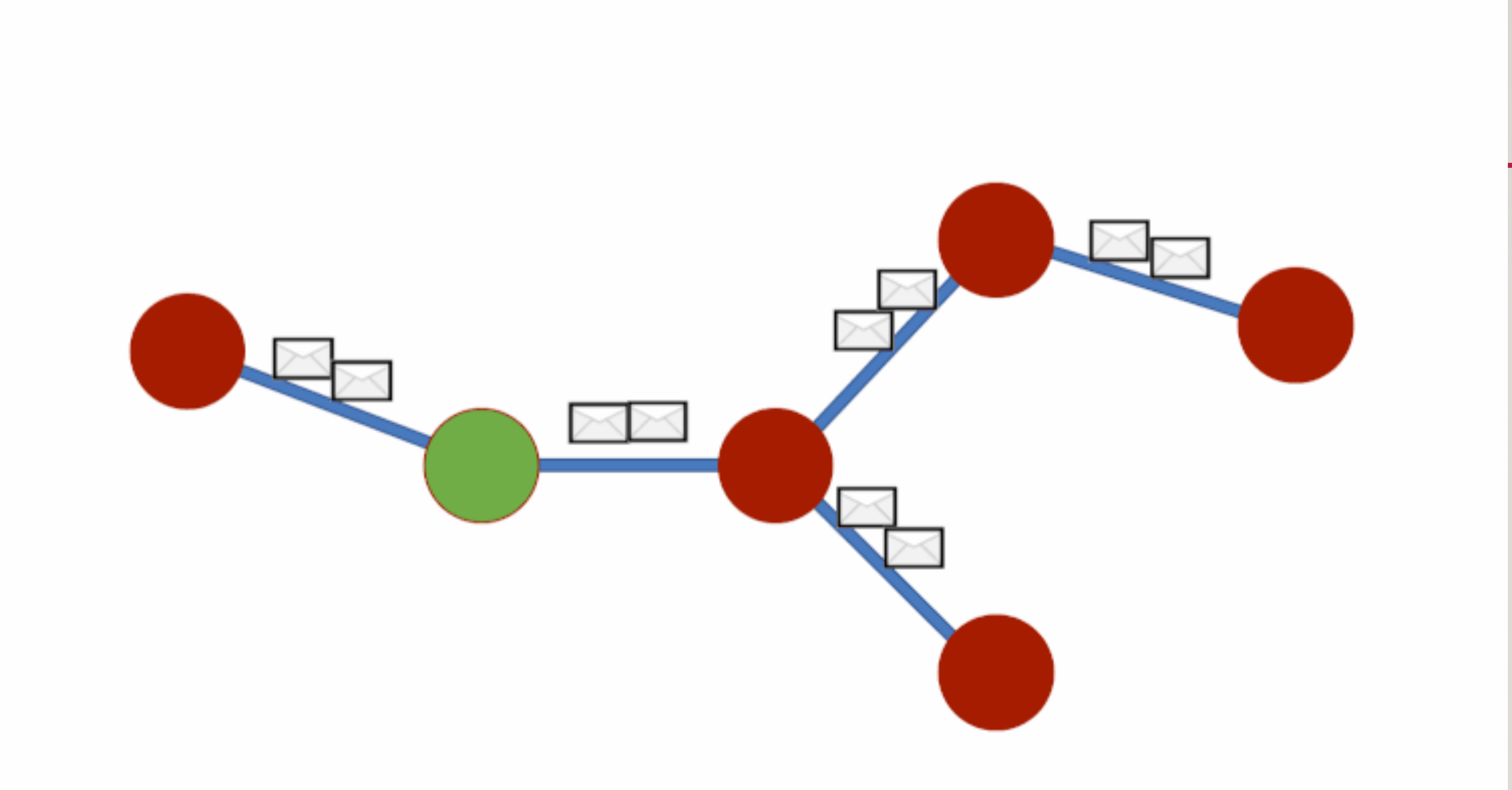


prasannavj@gmail.com



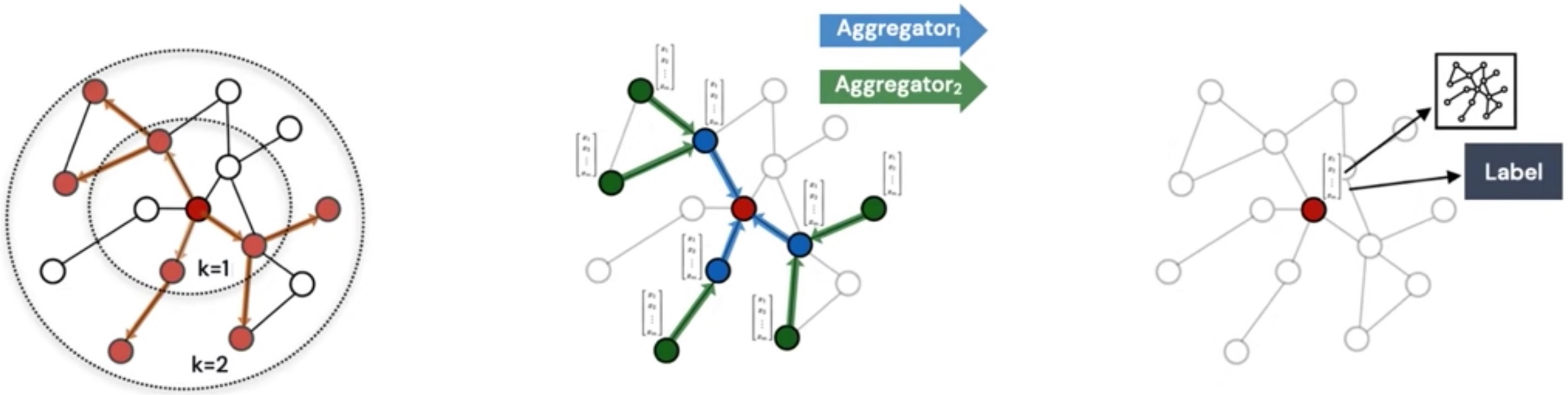
prasannavj@gmail.com

Message Passing Graph Neural Network



GraphSAGE (Hamilton et al., NeurIPS 2017)

How would you run a GNN on **very large** graphs?

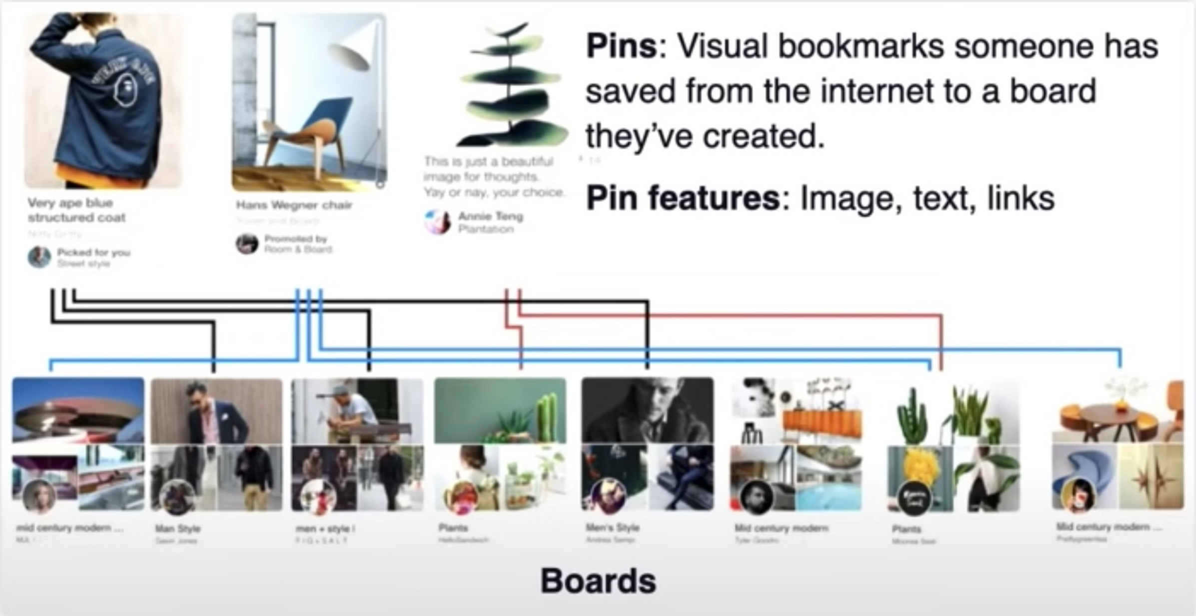


1. Sample neighborhood
2. Aggregate feature information from neighbors
3. Predict graph context and label using aggregated information

prasannavj@gmail.com

PinSAGE (Ying et al., KDD 2018)

Taking GraphSAGE to the extreme: apply to *Pinterest* graph (**3bn** nodes



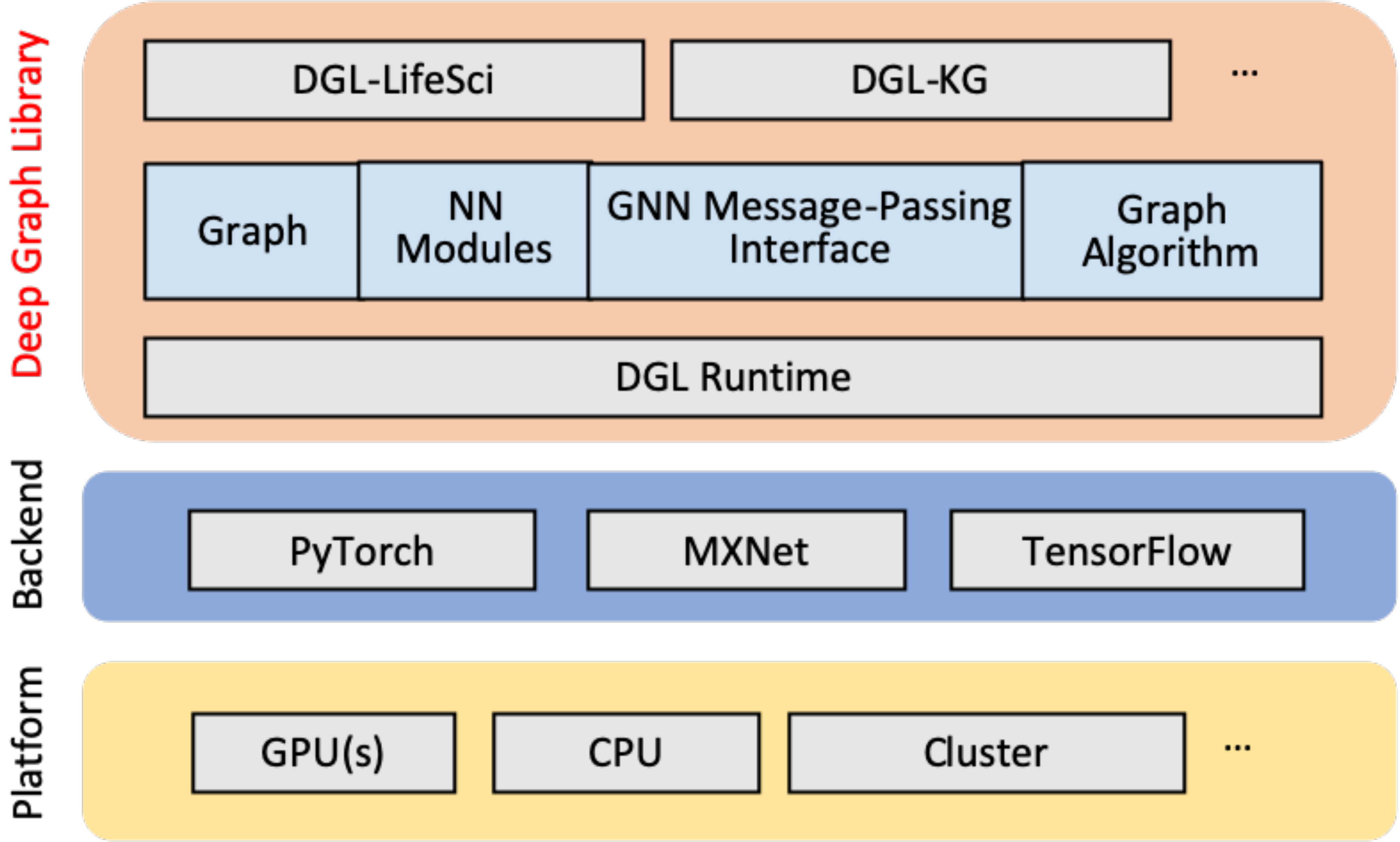
Use existing (pin, board) links in a **link prediction** task!

prasannavj@gmail.com

DGL – DEEP GRAPH LIBRARY -INTRO

- Open Source library for deep graph based machine learning operations.
- Promoted by AWS
- Support Python

prasannavj@gmail.com


Deep Graph Library
DGL-LifeSci
DGL-KG
...
Graph
NN Modules
GNN Message-Passing Interface
Graph Algorithm
DGL Runtime
Backend
PyTorch
MXNet
TensorFlow
Platform
GPU(s)
CPU
Cluster
...

DGL – LIVE DEMO

https://github.com/prasannavj/GraphMachineLearning