Webinar on Analytics and Data Science

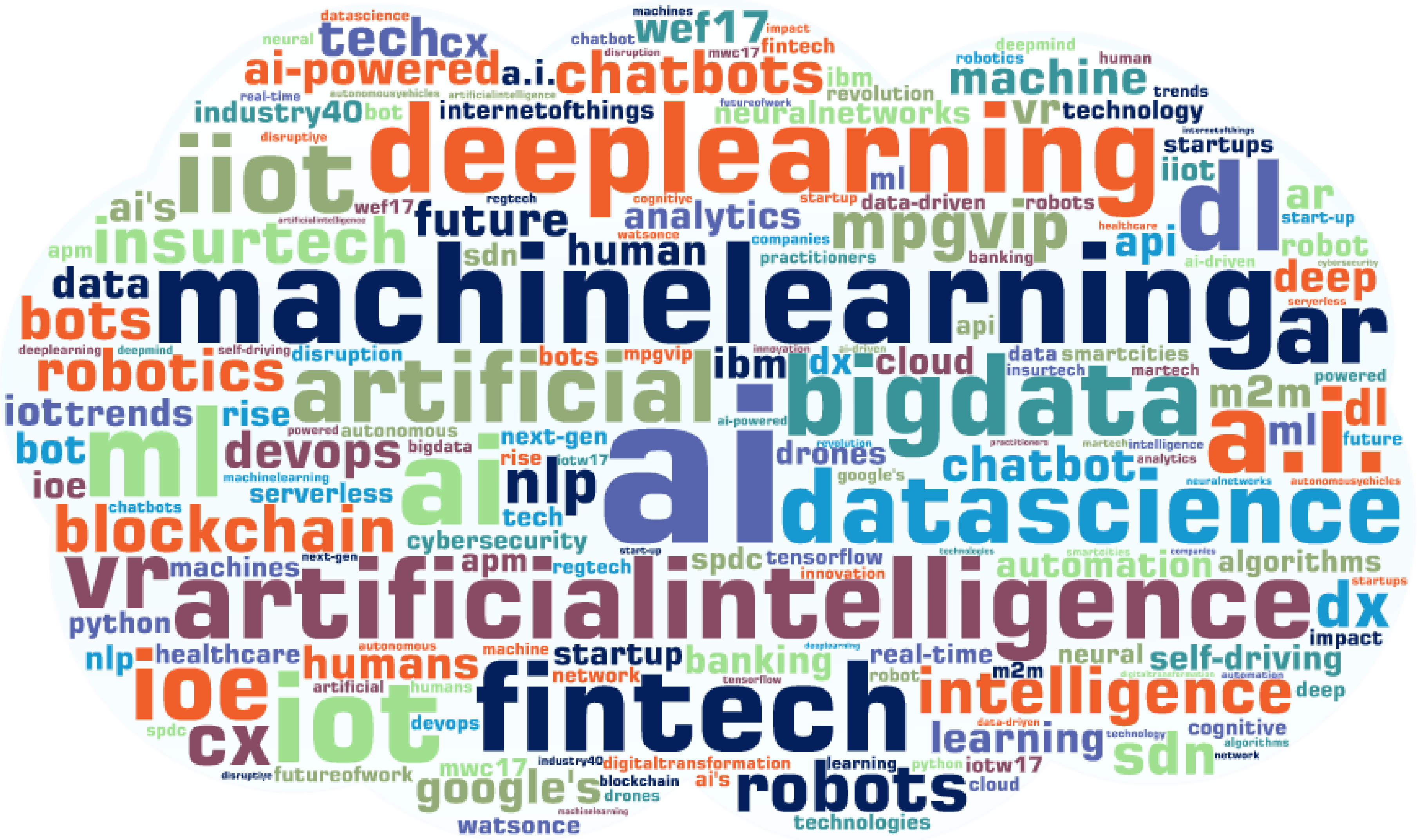
11-June-2022

AI in Fraud and Anomaly Detection

Santosh Bhat
Head of Data Sciences

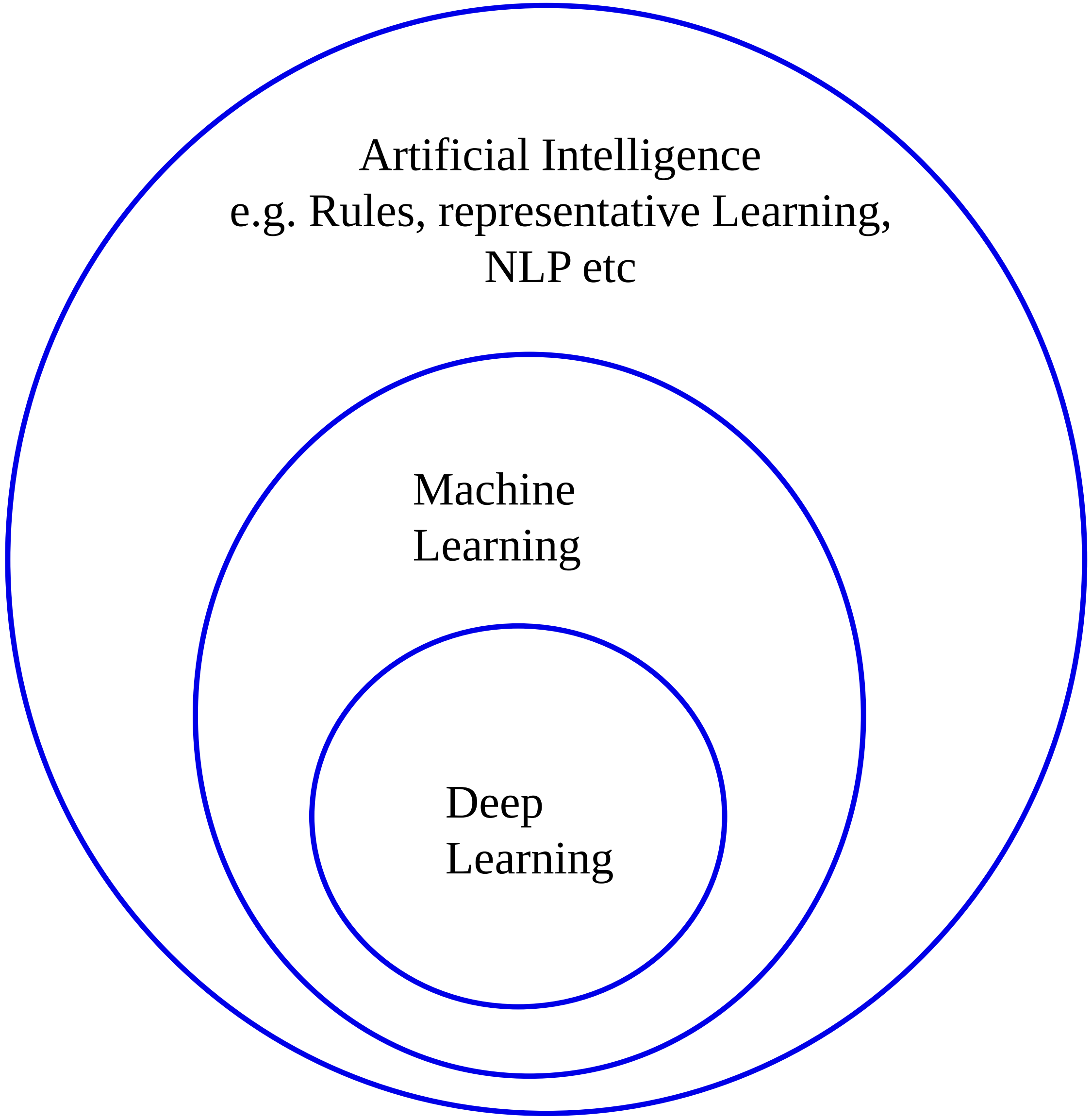
Institute of Actuaries of India

Buzzwords!!



Artificial Intelligence is here to stay!!

Let’s start with some definitions!!



What is the difference between AI and Machine Learning?

Artificial intelligence (AI):
Building smart machines capable of performing tasks that typically require human intelligence.

Machine Learning:
Machines or Algorithms learning through data without having to explicitly program the same

Deep Learning – A field of Machine Learning that uses multiple layers of Neural Networks to perform specific tasks

Machine Learning is considered to be the subset of AI

AI applications in Insurance

Institute of Actuaries of India

Chatbots

Voicebots

Speech to Text

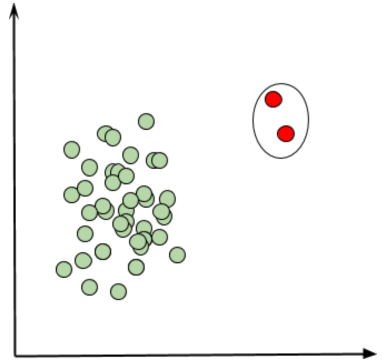
Computer vision

Natural Language processing

Propensity/Churn Models

Fraud Models

Anomaly detection

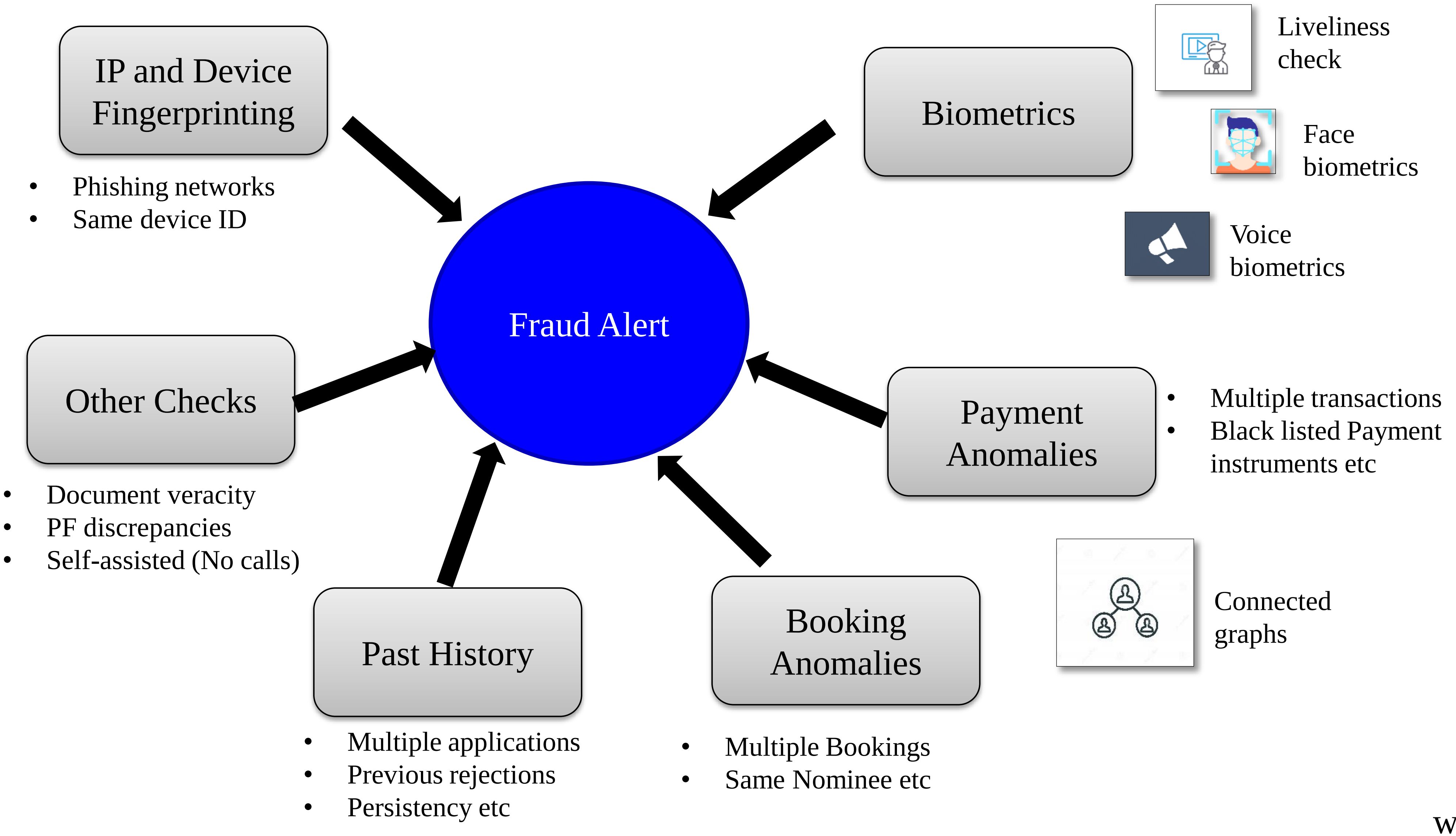


Personalized recommendations

Marketing Analytics

www.actuariesindia.org

Setting up a comprehensive Fraud framework is important!!



Fraud Alert

IP and Device Fingerprinting

- Phishing networks
- Same device ID

Biometrics

- Liveliness check
- Face biometrics
- Voice biometrics

Other Checks

- Document veracity
- PF discrepancies
- Self-assisted (No calls)

Payment Anomalies

- Multiple transactions
- Black listed Payment instruments etc

Past History

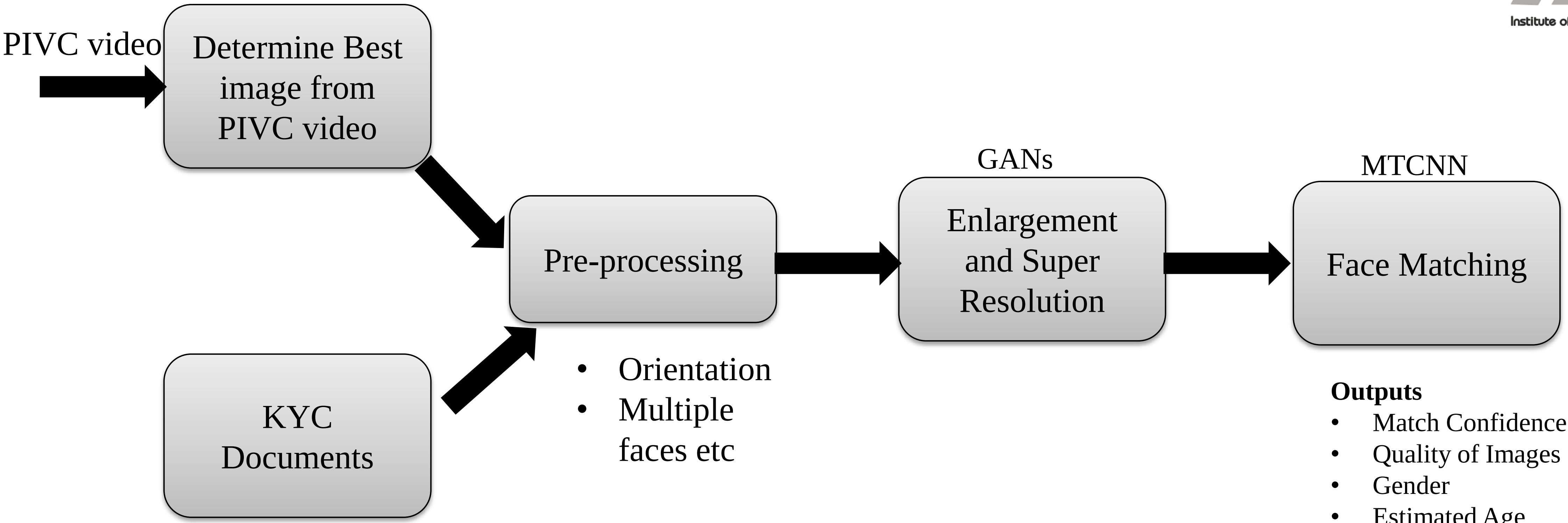
- Multiple applications
- Previous rejections
- Persistency etc

Booking Anomalies

- Multiple Bookings
- Same Nominee etc

Connected graphs

Face Biometrics



PIVC video

Determine Best image from PIVC video

KYC Documents

Pre-processing

GANs

Enlargement and Super Resolution

MTCNN

Face Matching

- Orientation
- Multiple faces etc

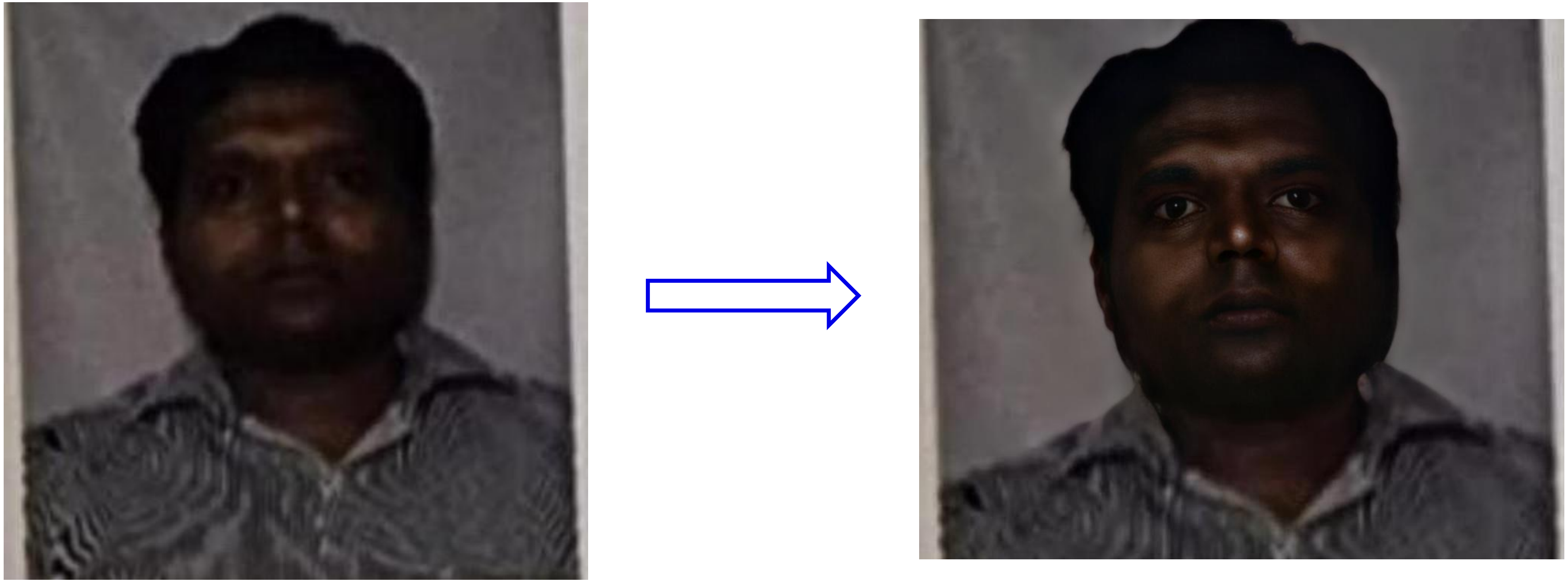
- Aadhar
- Pancard
- DL
- Voter ID
- Passport etc

Outputs

- Match Confidence
- Quality of Images
- Gender
- Estimated Age
- Gear (Turban, glasses, Masks)

In Development:
Background prediction (Temp hut, park etc)

Super Resolution of Images



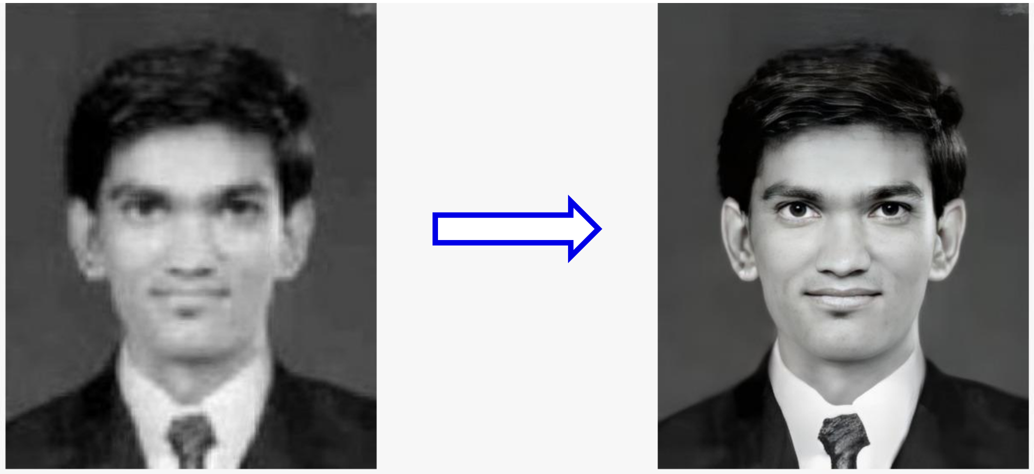
- Generative Adversarial Networks (GAN) based algorithms help in enhancing resolution of images, particularly of KYC documents

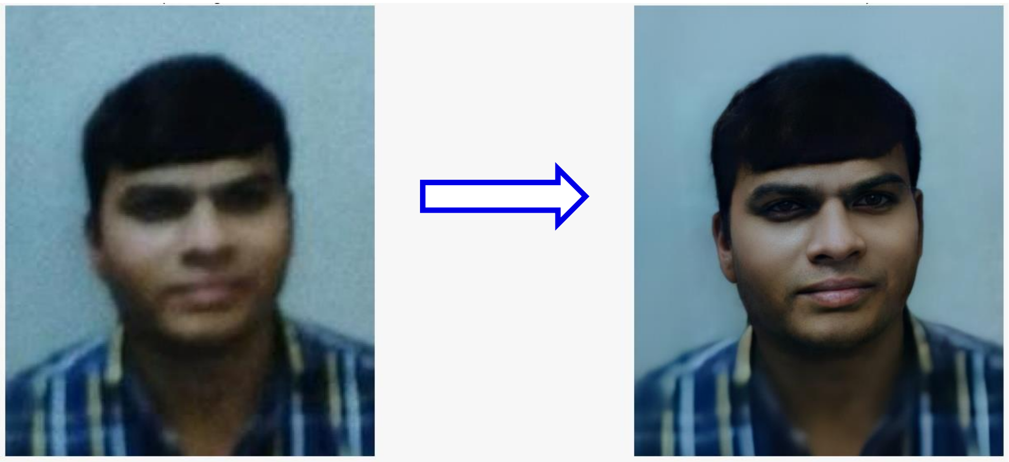
- Helps improving the confidence of matches in the Facial recognition phase
  - Also helps verification teams in verification

Image Augmentation using GANs

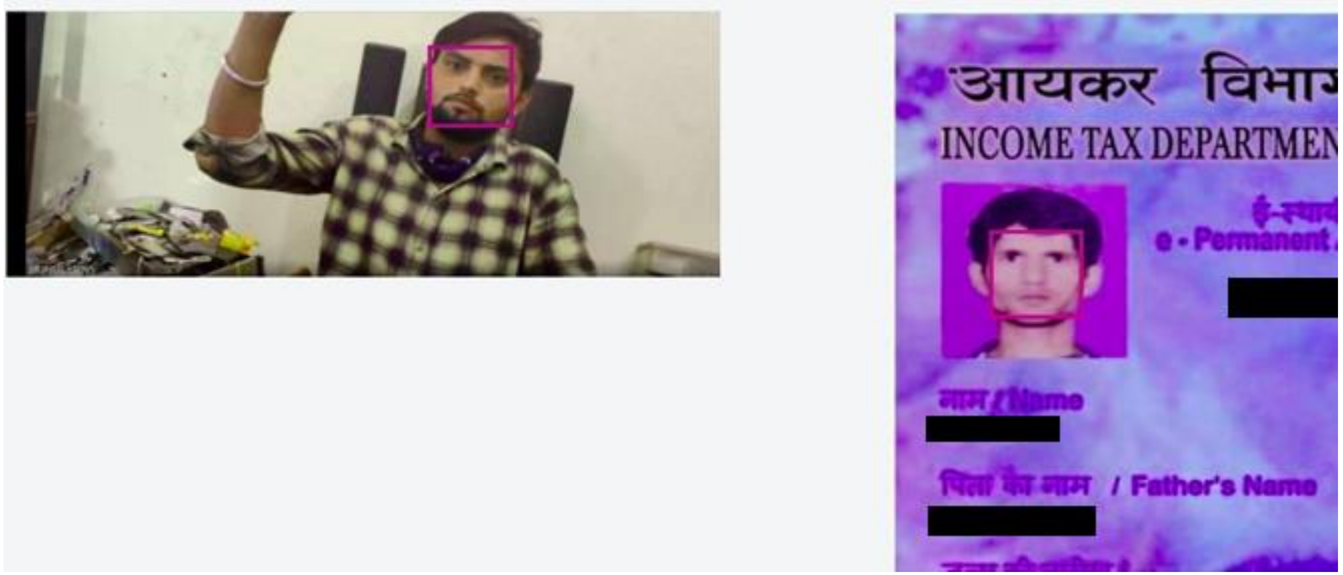
[redacted]

[redacted]

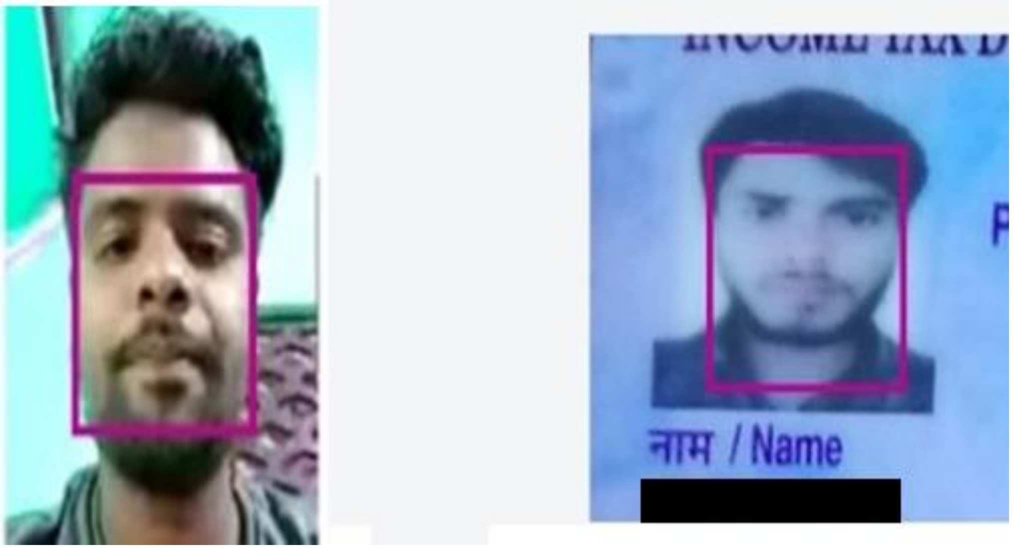




Facial Recognition detects Impersonation

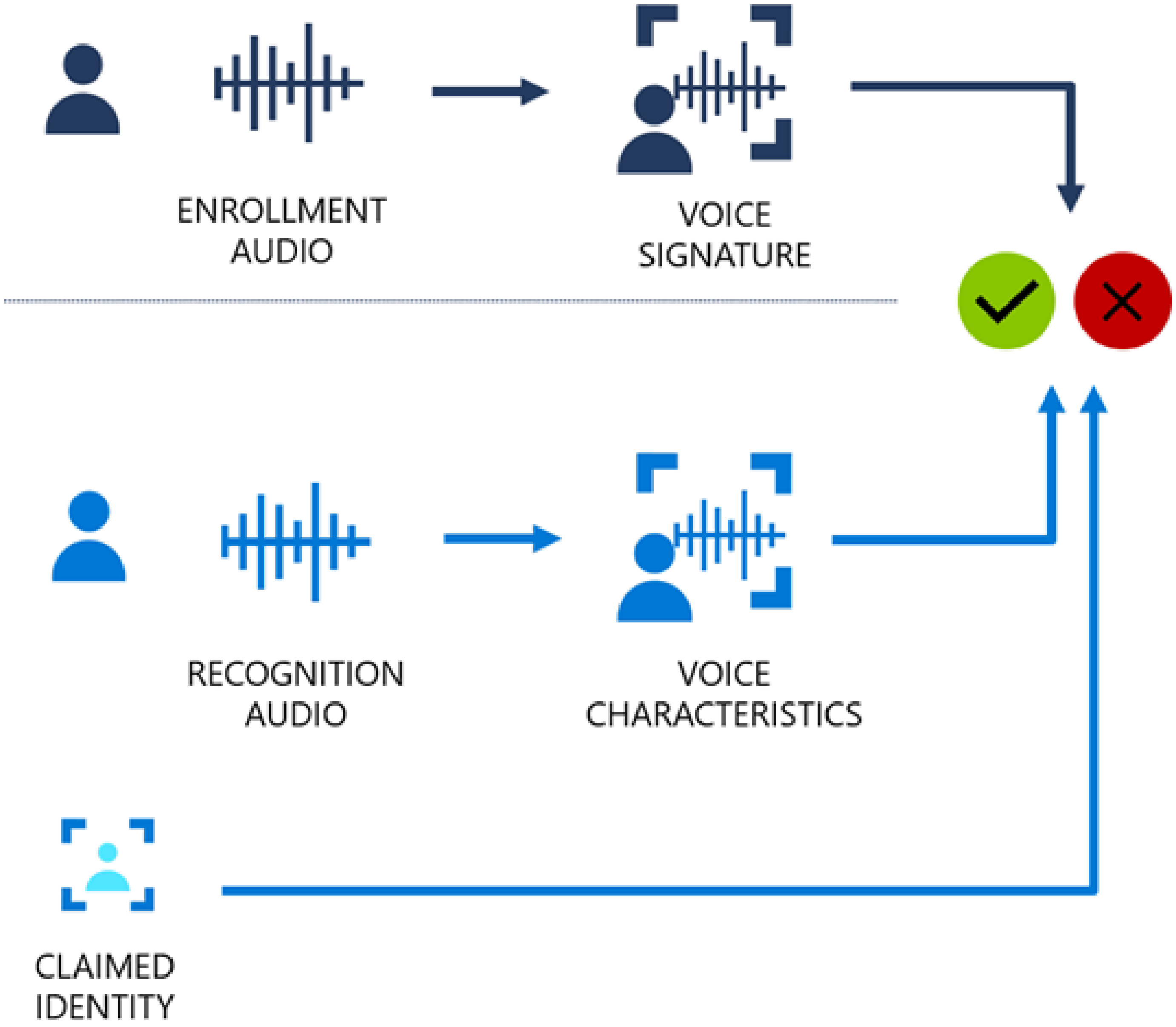


Confidence of Match: 0.0905. Two different people



Confidence of Match: 0.10109. Two different people

Voice Verification



Benefits:

1. Making sure that the person across both the calls are the same
2. Can be used as a speech biometric for verification in case a claim is raised

Methodology:

Works on Deep Learning model that compares two spectrograms.

Several innovations incorporated:

- Background Noise reduction
- Blank segment removal
- Specific segments of customer audio in training
- Removal of Ringtones and other noises from training audio
- Gender detection

Payment Anomalies

Types of Payment Anomalies

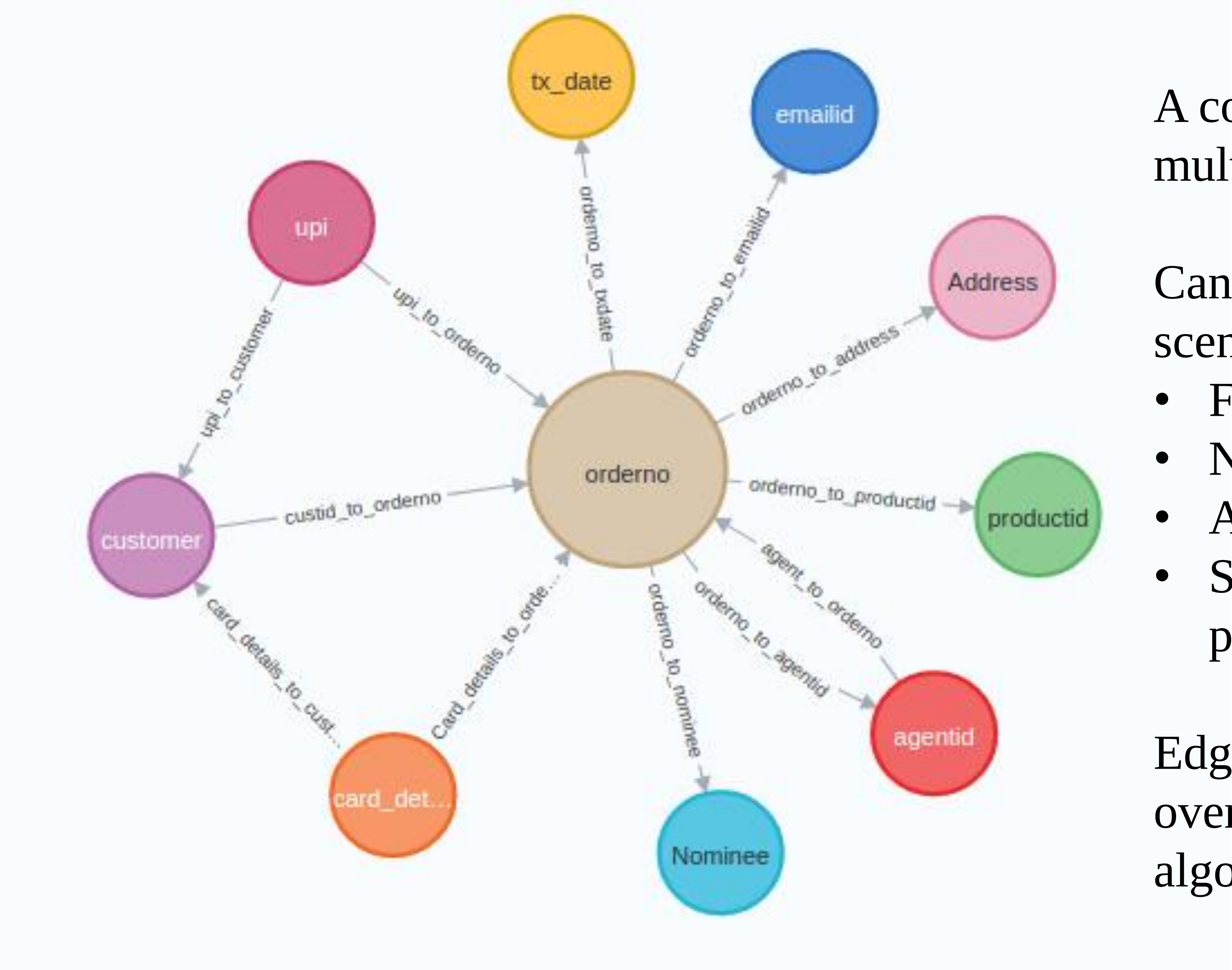
- One card Multiple Payments
- One card multiple customers
- Too many payments within a short time span
- One Person – Multiple Payment Instruments (UPI, CC, DC)
- Brokers from different Business units (GI)

Booking Anomalies

- One nominee Multiple proposers
- Same email ID being used by multiple customers
- Same Physical address shared by multiple customers
- Fraud Rings – multiple groups working together

How do we develop a framework to detect “Interesting” patterns?

Graph based Anomaly detection



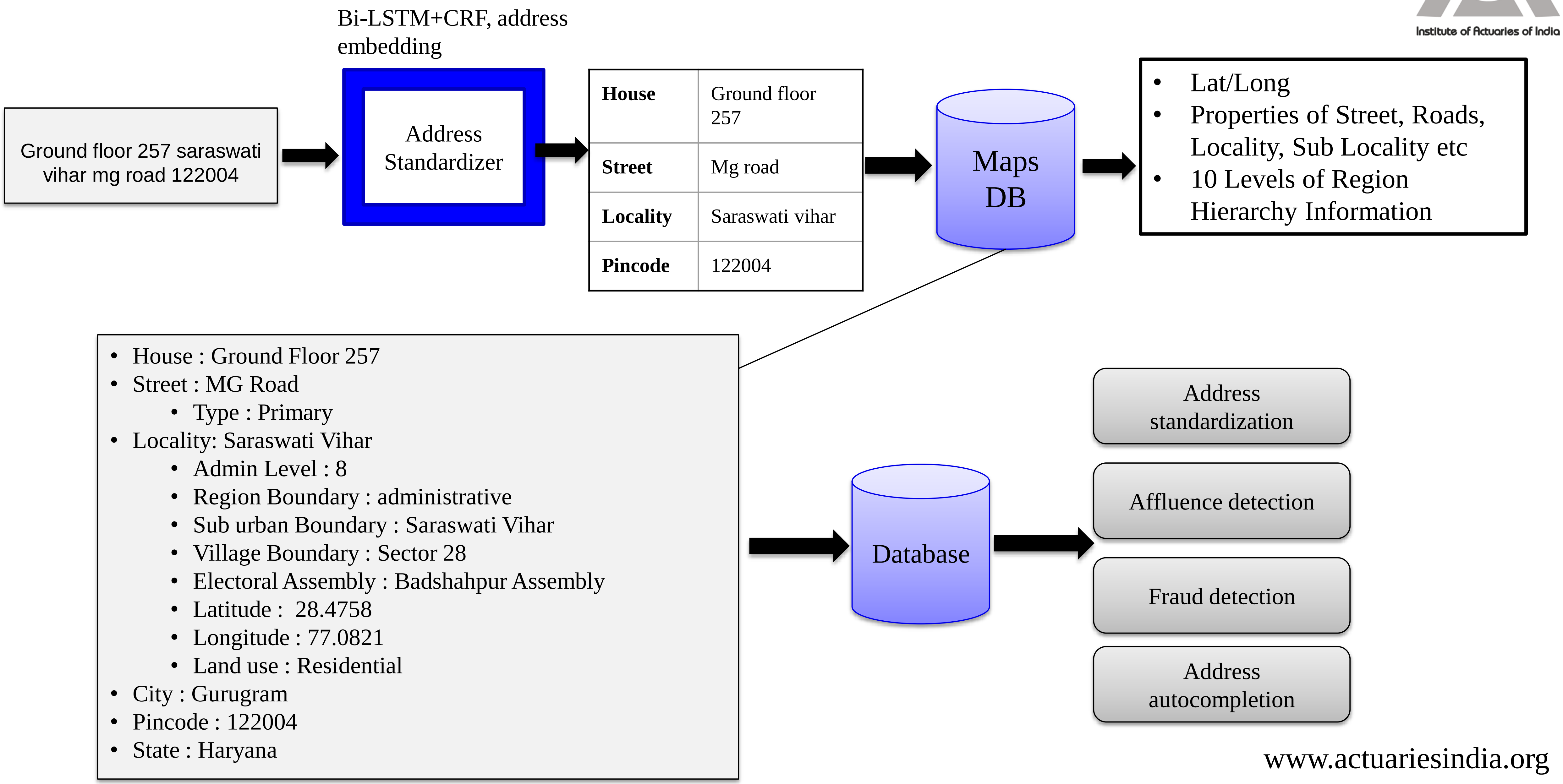
A connected graph can help combine multiple attributes of a customer

Can help uncover complex fraud scenarios

- Fraud rings
- Nexus between Agent – Customer
- Agent related Frauds
- Suspicious transactions basis payment instruments etc..

Edges of the graph can be weighted over a period of time with a ML algorithm

Addresses Intelligence



Ground floor 257 saraswati vihar mg road 122004

Bi-LSTM+CRF, address embedding

Address Standardizer

House	Ground floor 257
Street	Mg road
Locality	Saraswati vihar
Pincode	122004

Maps DB

- Lat/Long
- Properties of Street, Roads, Locality, Sub Locality etc
- 10 Levels of Region Hierarchy Information

- House : Ground Floor 257
- Street : MG Road
    - Type : Primary
- Locality: Saraswati Vihar
    - Admin Level : 8
    - Region Boundary : administrative
    - Sub urban Boundary : Saraswati Vihar
    - Village Boundary : Sector 28
    - Electoral Assembly : Badshahpur Assembly
    - Latitude : 28.4758
    - Longitude : 77.0821
    - Land use : Residential
- City : Gurugram
- Pincode : 122004
- State : Haryana

Database

Address standardization

Affluence detection

Fraud detection

Address autocompletion

Summary

- AI and ML technologies have come of age in solving fairly complex problems in every domain.
  - Not all problems can be solved with AI though!!

- Fraudsters are always ahead of the game. Technology needs to evolve at a rapid pace to keep up with Fraudsters

- Pay attention to the data that is available. More Data >> Better Algorithms

- “Human-in-the-loop” AI based frameworks work well to cover for any biases that the Algorithms may bring in

- Always keep asking Questions. Manual testing will help uncover a lot of use cases of fraud that any technology may not be able to solve!

THANK YOU