Machine Learning with Python & AI Foundation Course

May, 2018

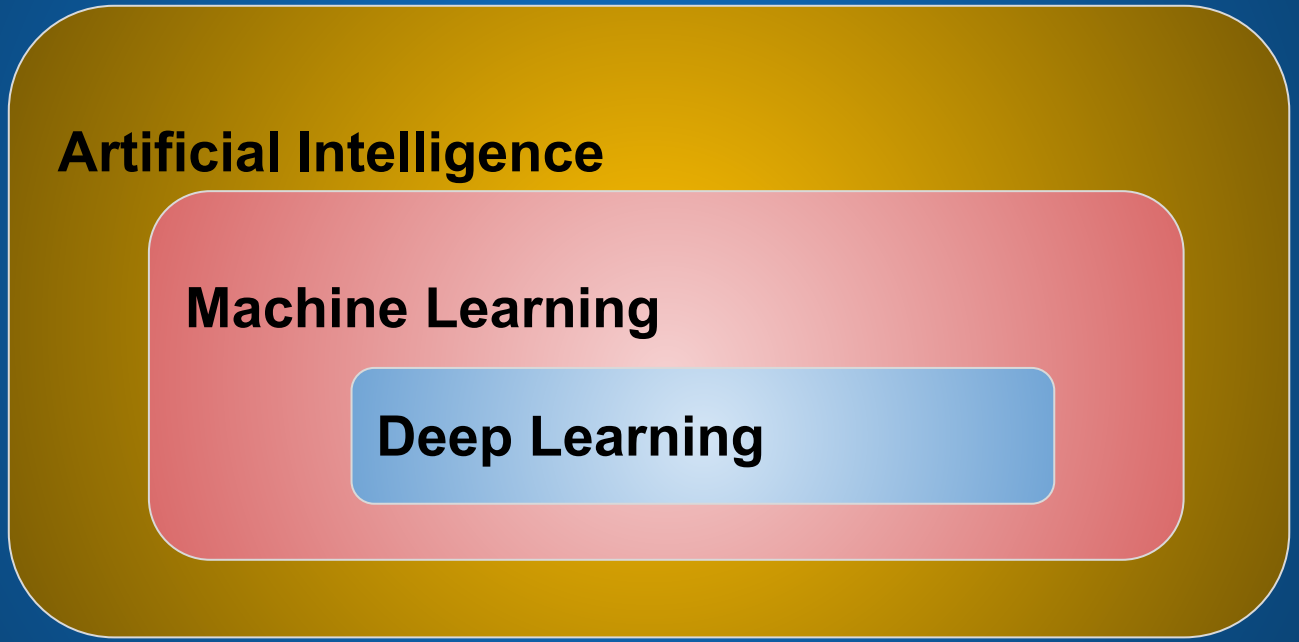
Course Structure

- Generic Programming Structure
- Python Programming
- Jupyter Notebook
- Regression, Classification, Clustering
- Association Rule Learning
- Reinforcement Learning
- Introduction to NLP & Deep Learning
- Dimensionality Reduction
- Model Selection & Boosting

Lecture One.

AGENDA

- *Introduction to AI and ML.*
- *General Programming Structure*
- *Python Libraries*
- *Machine Learning Principles*


Artificial Intelligence
Machine Learning
Deep Learning

What is Artificial Intelligence?

- An artificial intelligent system is one that can *learn* on its own.
- Can Improve on Past Iterations.
- Get Smarter with Time.

Is AI a Threat to Human Society?

- **No!**
- AI will not replace humans.
- The purpose of AI is to augment and simplify the purpose of complex tasks.
- Invention of wheel Analogy.

Applications of Artificial Intelligence

- Siri/Cortana/Iris *(NLP and Speech Recognition)*
- Alexa/Google Assistant *(IOT)*
- VR and AR *(Computer Vision)*
- Self Driving Cars *(Deep Learning)*
- Netflix Recommendation Engine *(Predictive analytics)*

Machine Learning

“ An Algorithm that learns from dynamic datasets, identifies patterns within large corpus of data and generates a model that accurately depicts the behaviours of the processes interacting in the system”

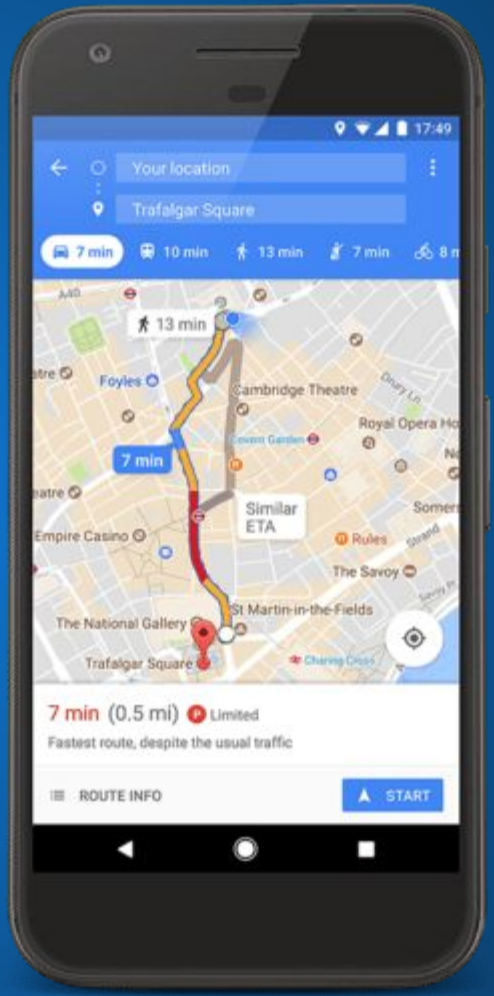
Commercial Value of Machine Learning

- Machine Learning models help us generate "***actionable insights***" of great value to the customer, client or society as a whole.
- "***Precise***" future predictions of engineered processes using data gathered through sensors, satellites, legacy databases etc.
- Predicting Machine Failure using sensor data. Actionable Insight? "***Preventive Repair.***"
- Industry Sector: Healthcare, Automobile, Factories, Quality Service, Maintenance etc.

Machine Learning in Everyday Life

Google Maps

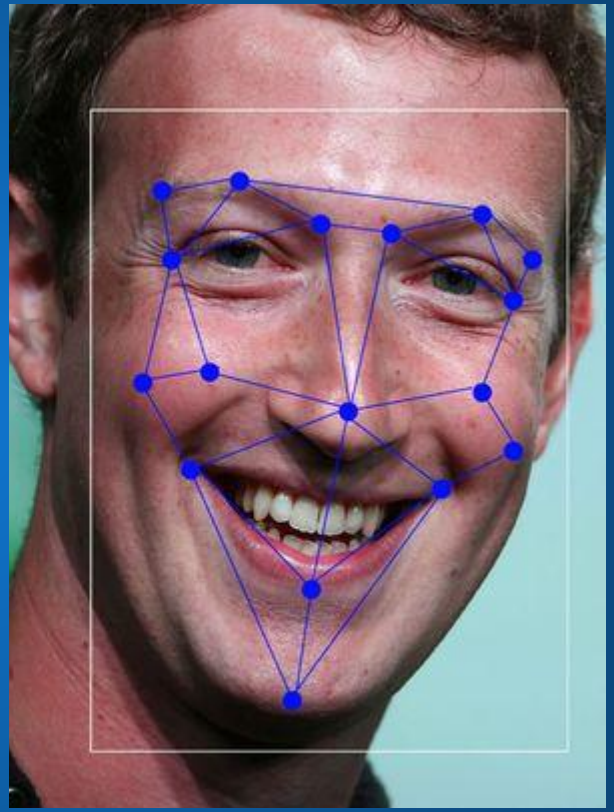
- ➔ Navigation
- ➔ Traffic Updates
- ➔ Street view



DeepFace

Facebook's
Face Recognition
Algorithm

➔ Based on RNN

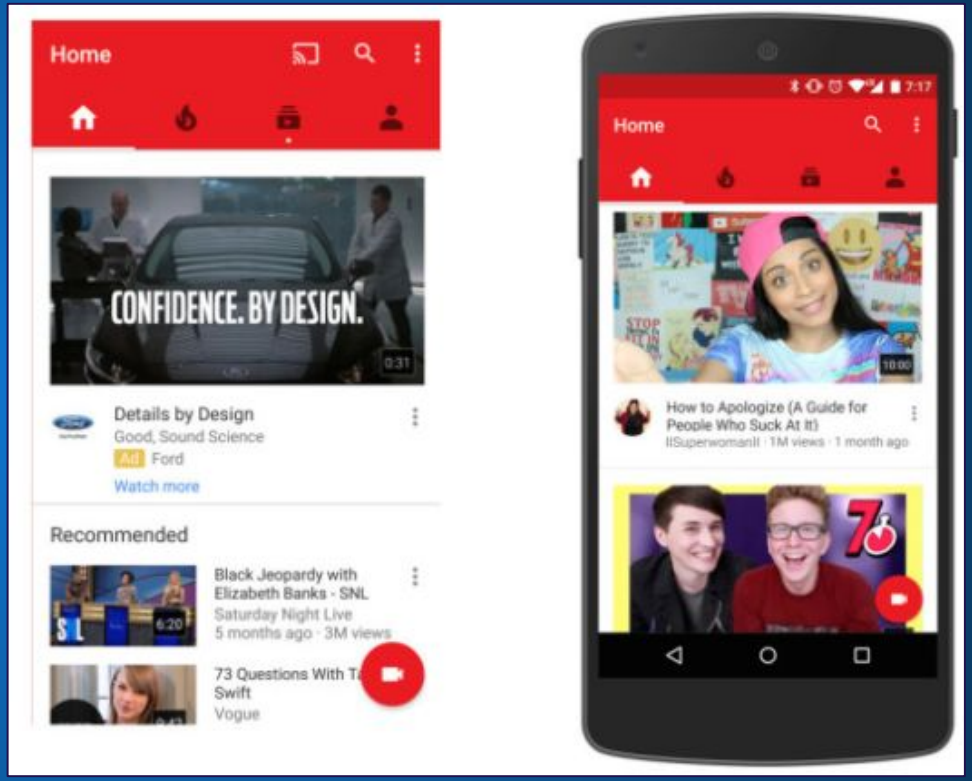


Uber Rides

- ETA
- Identifying exact pick up points and STOP offs.
- Predicting Tariff based on Road Conditions.

Youtube

- Automatically Generated Playlists
- Customized/ Personalized Ads
- Recommending better videos
- Improving Home Experience



Amazon

- Product Suggestions
- Virtual Assistant Echo & Dot
- Alexa
- Speech and Language

Why Python?

- ★ Open Source Software
- ★ Supports many important Machine Learning libraries and packages.
- ★ Easy to understand and follow.
- ★ Fast Learning Curve
- ★ Industry wide reach
- ★ Great for Beginners!

Python ML Libraries

- NumPy
- SciPy
- Pandas
- Scikit
- Tensorflow
- Keras
- matplotlib

NumPy

"Fundamental package for all scientific computations and statistical fxns. Also effective for utilizing generic data in multidimensional containers."

SciPy

“ The SciPy library is one of the core packages used for applying ML techniques. It provides many user friendly and efficient numerical routines used for numerical integrations, differential equations, optimization etc.”

matplotlib

“ Matplotlib is a plotting library for python programming language. It generates high quality figures such as histograms, bar charts, scatter plots etc. It is the single most used package for 2D graphics in python. ”

Scikit

“ Efficient tool for data mining and data analysis such as Clustering, Regression and Classification. Built on Numpy, SciPy and matplotlib. Great for exploratory analysis. Scikit stands for Scipy Toolkit.”

Keras

“ Keras is a high level neural networks API, written in Python. It supports both Convolution Neural Network (CNN) and Recurrent Networks (RNN) etc. Efficient in image processing tasks.”

Pandas

“ A collection of Data Structures and Tools for Data Analysis. Widely used in finance and engineering sectors. Works best with messy, incomplete, unlabeled real world data. Great collection of tools for reshaping, merging and slicing datasets.”

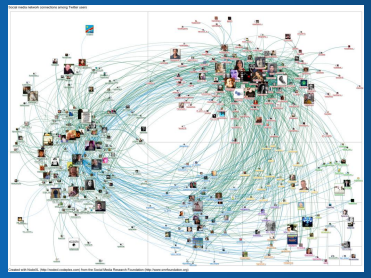
Tensorflow

“Developed by Google as an open source library. It is written in python and widely used for training neural networks.“

UNDERLYING PRINCIPLES BEHIND *MACHINE LEARNING*

Huge Volume of Data (Training Set: Features and Labels)

Assigning probabilities or weights to features

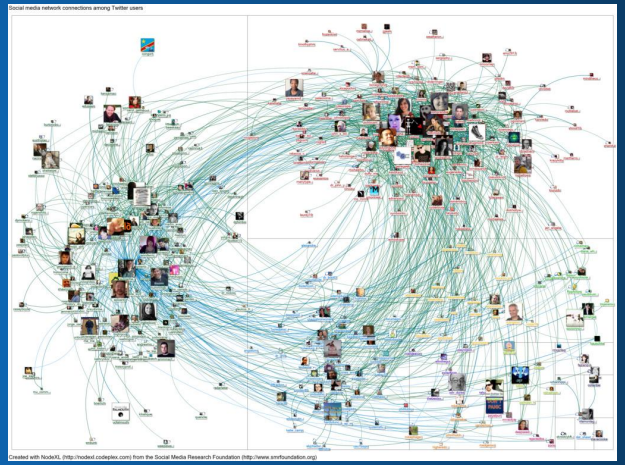


Generating a Preliminary Data model as closer to the true values as possible. This model will be build upon iteratively as more and more fresh data is plugged into the system.

UNDERLYING PRINCIPLES BEHIND *MACHINE LEARNING*

Huge Volume of Data (Terabytes of big data for more accurate representation)

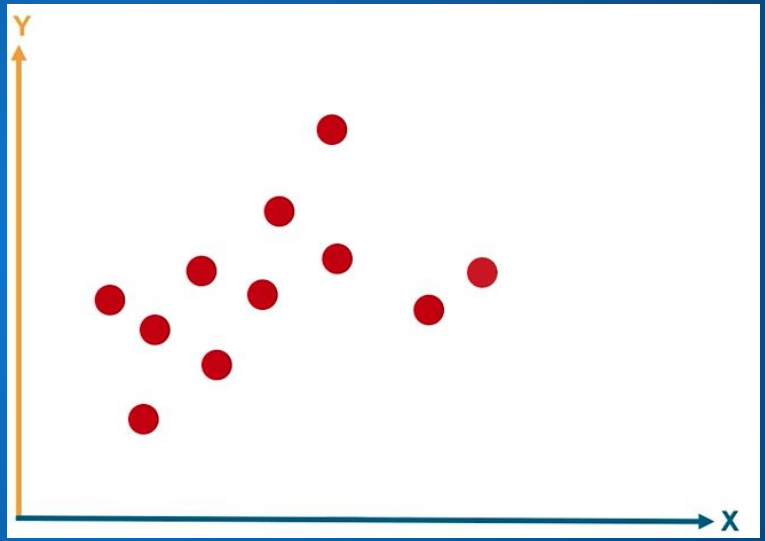
Mathematical or Statistical Function/Machine Learning Algo



A GDM uses nodes, edges, properties and relations to accurately predict the target values. Graph models are best to showcase highly interconnected data.

Machine Learning Example

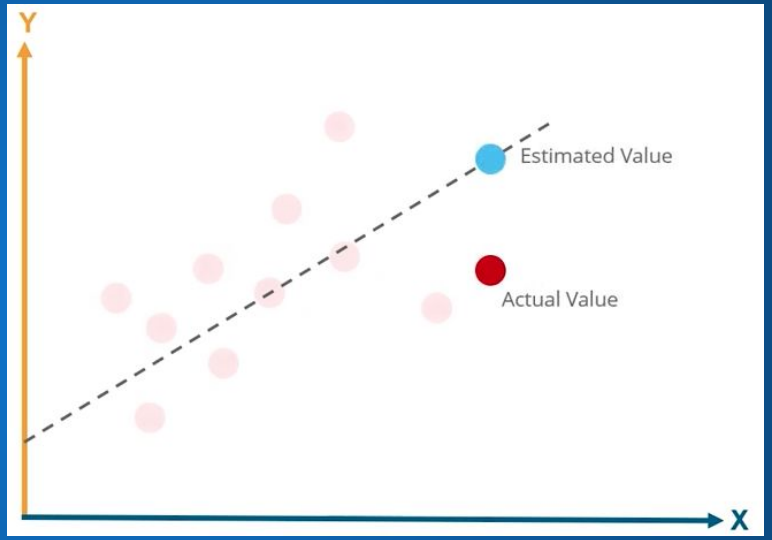
- ***Feature Variable :*** X
- ***Response Variable :*** Y
- ***X :*** Time
- ***Y :*** Temperature



Time Vs Temperature Plot

Machine Learning Example

- ➔ *F(x) = Y*
- ➔ *ax+b=y*



Time Vs Temperature Plot

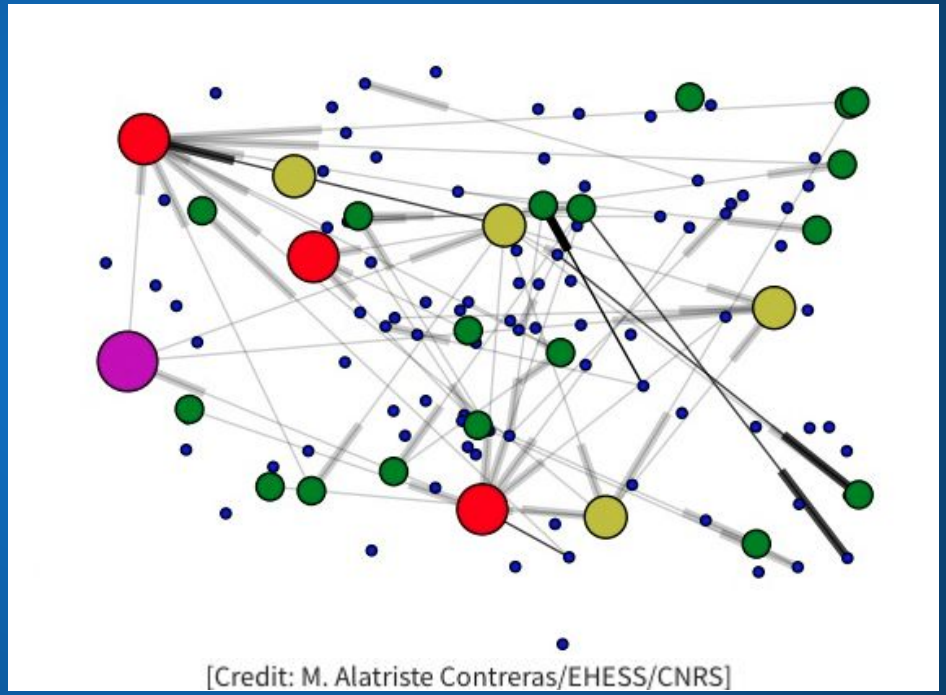
Machine Learning Example

- ***End Goal:*** Minimize the difference between Predicted Values and Actual Values
- Decrease in error increases performance of Model.

Time Vs Temperature Plot

CASE STUDY : International Trade Links

- ***Node :*** Country
- ***Edge :*** Trade Link
- ***Edge Strength :*** Amount of Trade
- ***Circles :*** Country
- ***Colour :*** Industrial Sectors
- ***Size of Circle:*** Net Production



[Credit: M. Alatriste Contreras/EHESS/CNRS]

CASE STUDY : International Trade Links

- Identify Business Problem

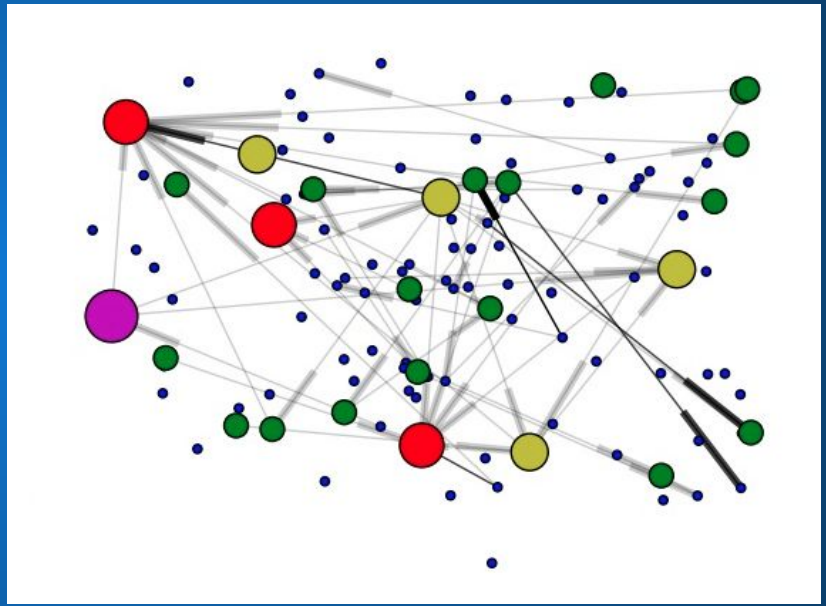
CASE STUDY : International Trade Links

- Potential Business Problem

“ When an economic shock occurs which country is likely to be affected the most? ”

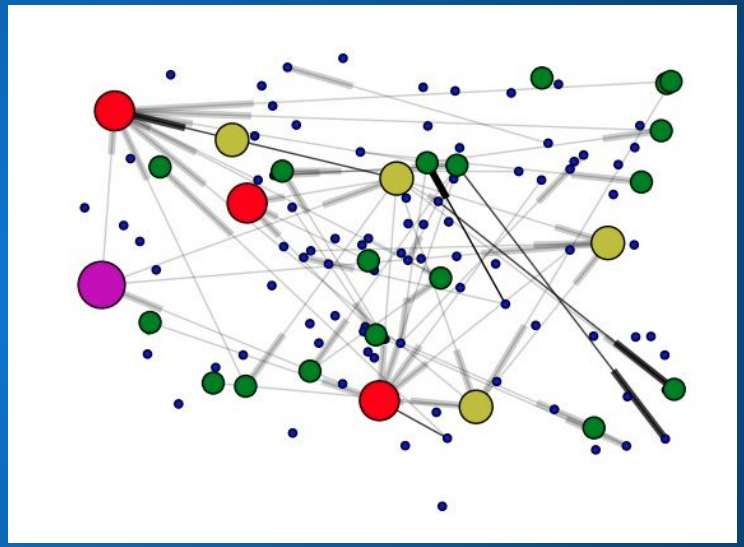
INSIGHTS : International Trade Links

➔ ***Countries with isolated trades would not spread shock as much.***



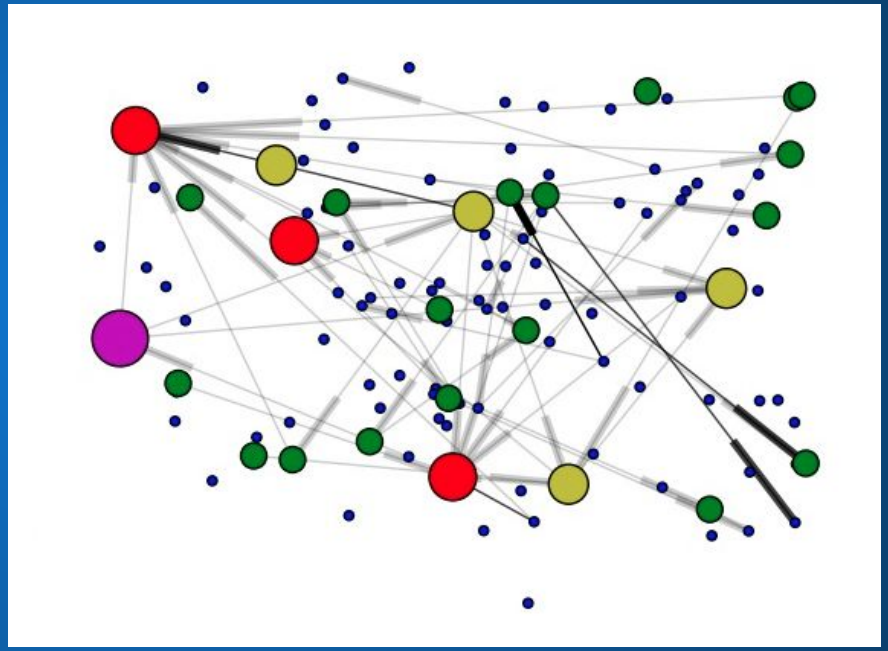
INSIGHTS : International Trade Links

- ***Countries which are more like industrial hubs having most number of direct and indirect connection spread the shock farthest and have the most potential to cause a global crisis.***



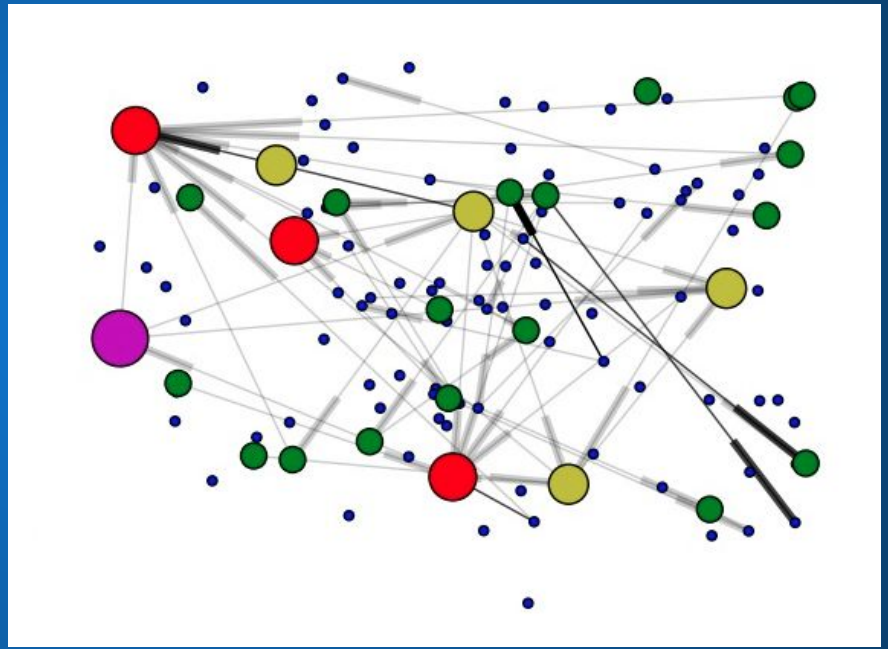
INSIGHTS : International Trade Links

- ***Countries with largest economies are most vulnerable because their trade links are highly developed and interconnected.***



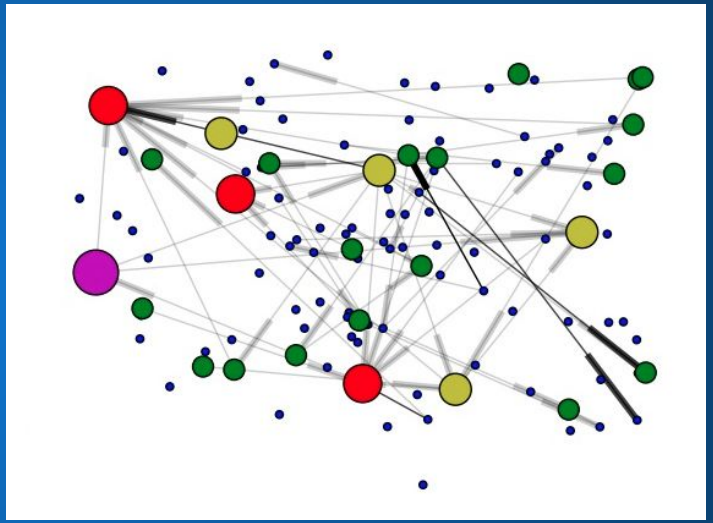
Quiz : International Trade Links

- ***Red = Category A***
- ***Purple = Category B***
- ***Olive = Category C***
- ***Green = Category D***
- ***Blue = Category E***



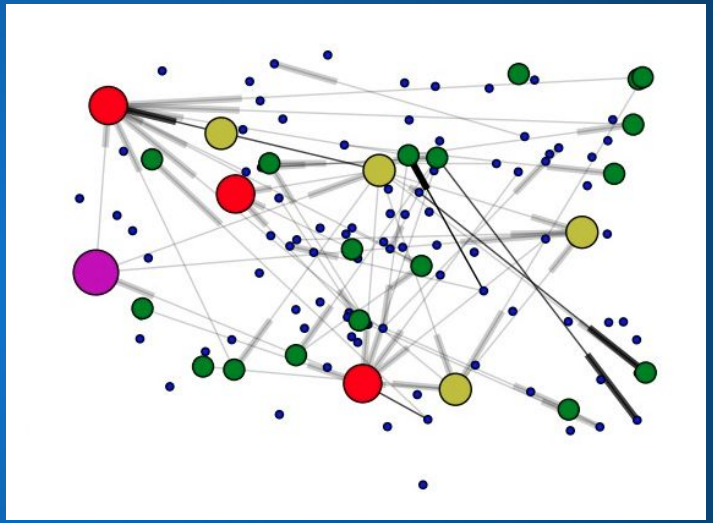
QUIZ 1 : International Trade Links

➔ ***Which Country will least spread a shock incase of an economic meltdown?***



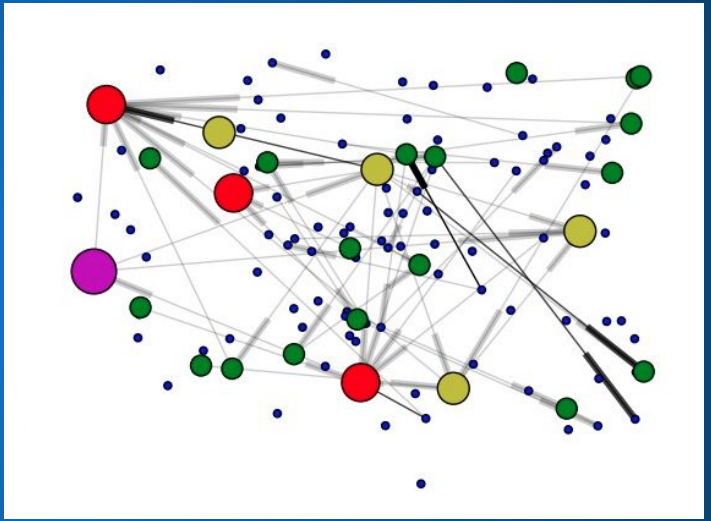
Answer 1 : International Trade Links

- ***Which Country will least spread a shock incase of an economic meltdown?***
  - ***Category E***



QUIZ 2 : International Trade Links

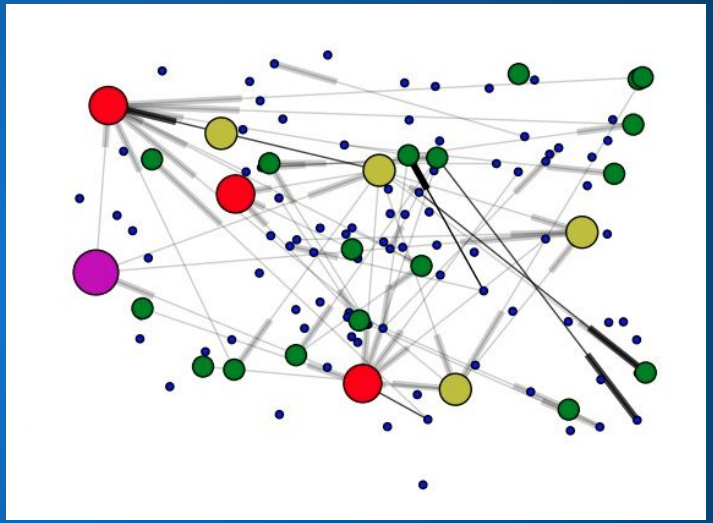
➔ ***Which category of countries have the highest potential to cause a global crisis?***



Answer 2 : International Trade Links

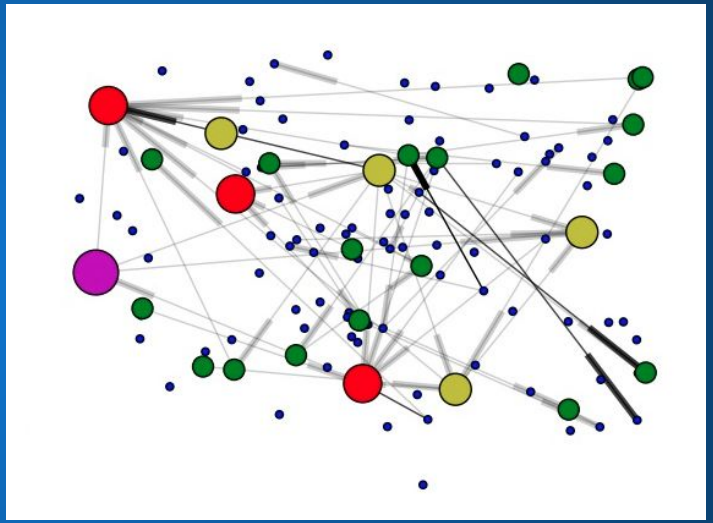
➔ ***Which category of countries have the highest potential to cause a global crisis?***

- ***Category D***



QUIZ 3 : International Trade Links

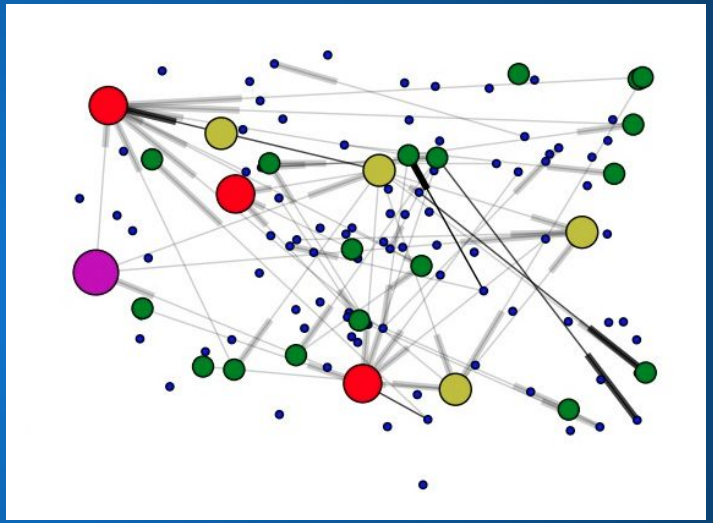
➔ ***Which category of countries is most vulnerable to global crisis?***



Answer 3 : International Trade Links

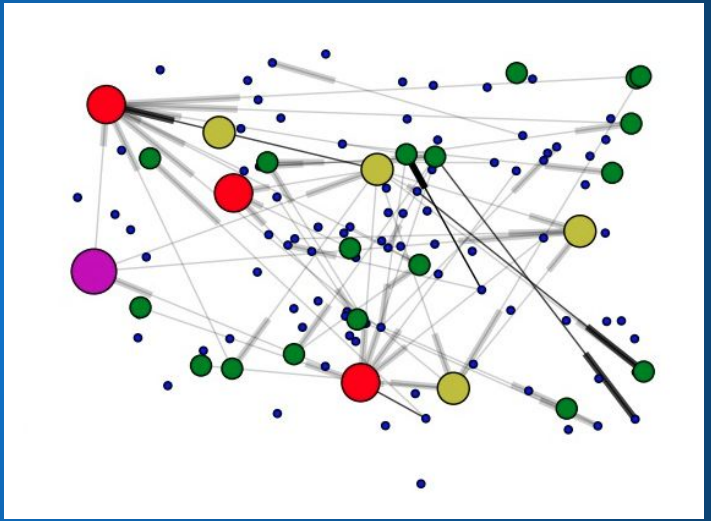
➔ ***Which category of countries is most vulnerable to global crisis?***

- ***Category A***



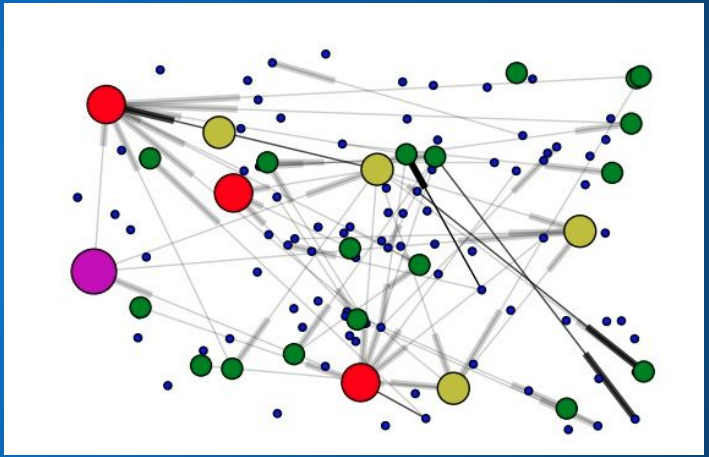
Quiz 4: International Trade Links

- ***Can you suggest other fields in which such graph data structure/model (GDM) could yield actionable insights?***



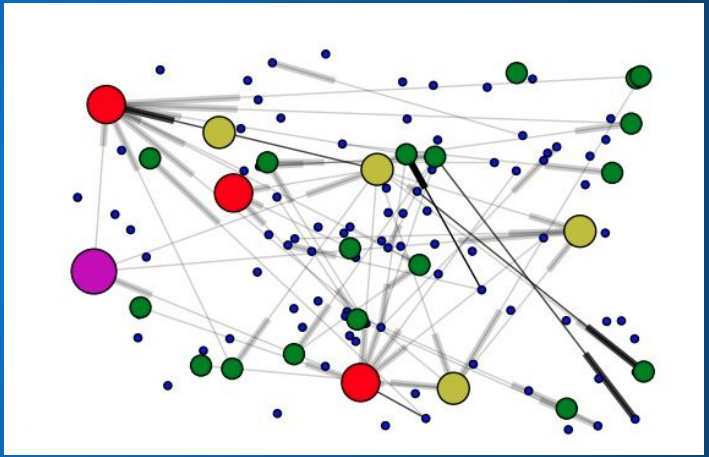
Answer 4 : International Trade Links

- ➔ *Fraud Detection*
- ➔ *Credit Risk Modelling*
- ➔ *Navigation Optimization*
- ➔ *Recommendation Engines*
- ➔ *Epidemic Prediction*



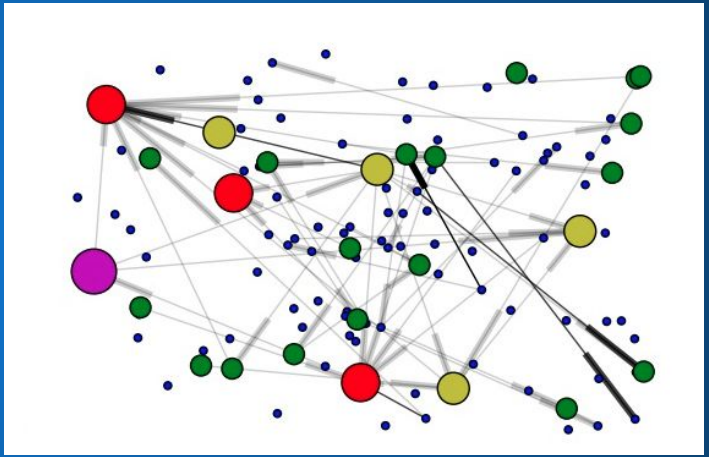
Answer 4 : International Trade Links

Business Problem: Which is the next most likely country to get affected by an ongoing epidemic. ”



Answer 4 : International Trade Links

“ What is the estimated box office revenue for the upcoming multi starrer movie ? ”



Summary

- *Machine learning is a subset of AI technique which uses statistical methods to enable machines to improve on themselves with experience.*

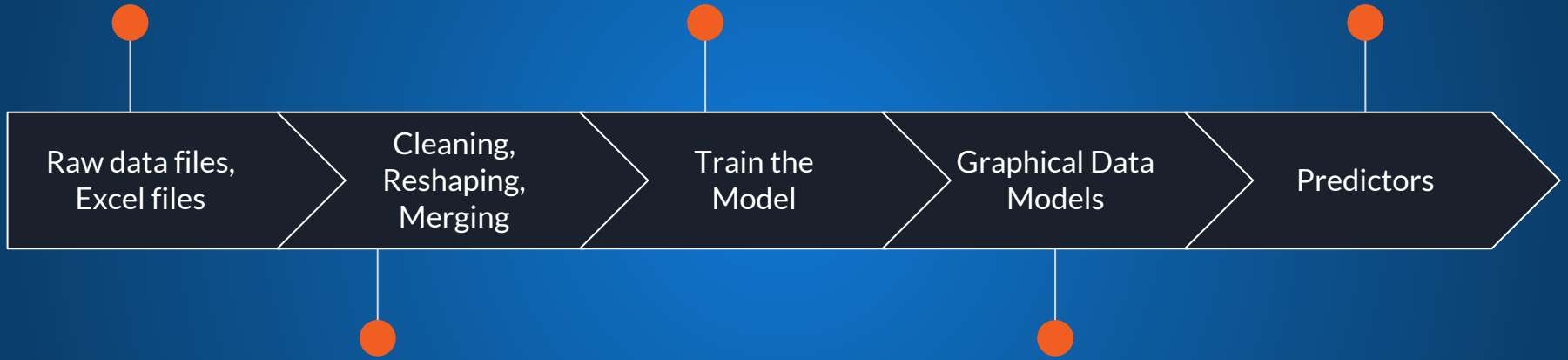
- *Machine Learning uses algorithms to parse data, learn from that data and make informed decisions based on what it has learned.*

Summary

Data Collection

Learning

Prediction/Analytics



Data Preparation

Data Modelling

- Python 2 Vs Python 3
- Python Vs JAVA/C

Assignment

- *Find out how Twitter uses AI and ML in improving customer experience?*

- *Further Reading :* https://blog.bufferapp.com/twitter-timeline-algorithm

End of Lecture One.