**Featurization & Feature Engineering:**

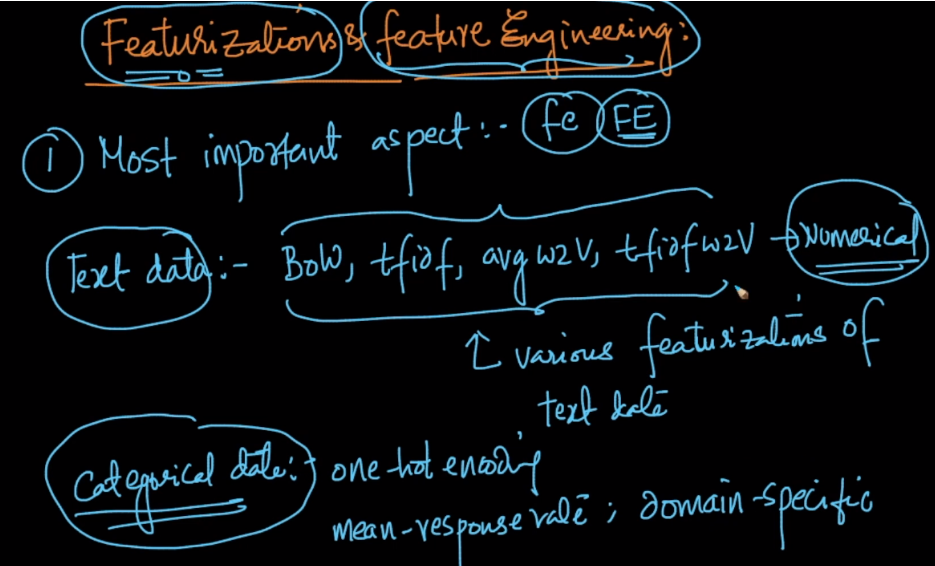
There are tons of feature engineering technique available, here we’ll look the techniques that are used usually.

1. **Text Data:**

There are various techniques we read like BOW, tfidf, avg w2v, tfidf w2v to convert text to numerical so that we can apply them to ML models.

1. **Categorical data:**

One hot enconding, and there can be others domain specific technique.



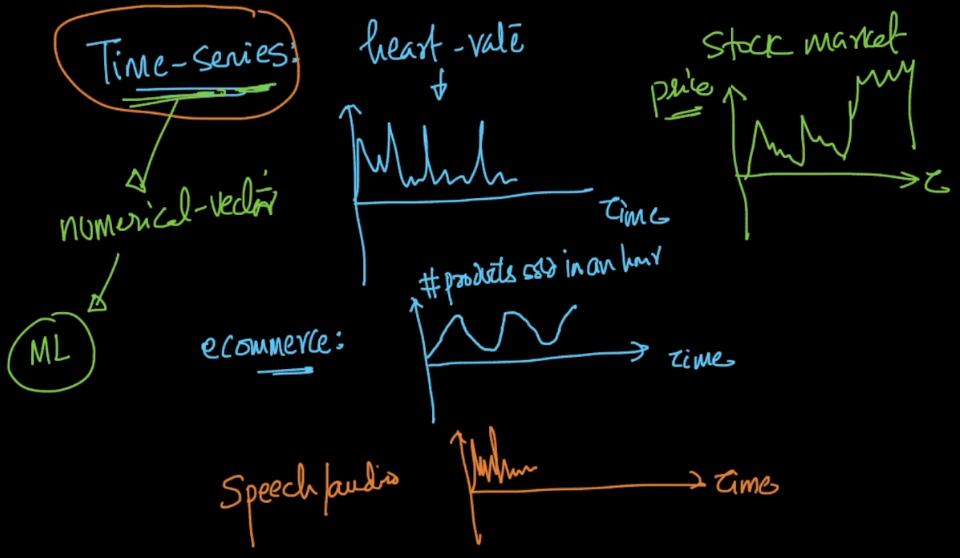
1. **Time-Series:**

Data which is related to time is said to be time series data, in this there is a sequence of data ie one data point occur only after that data point, because of time.

Example: stock market in which datpoint at 11 am should be come before datapoint at 11:15 am.

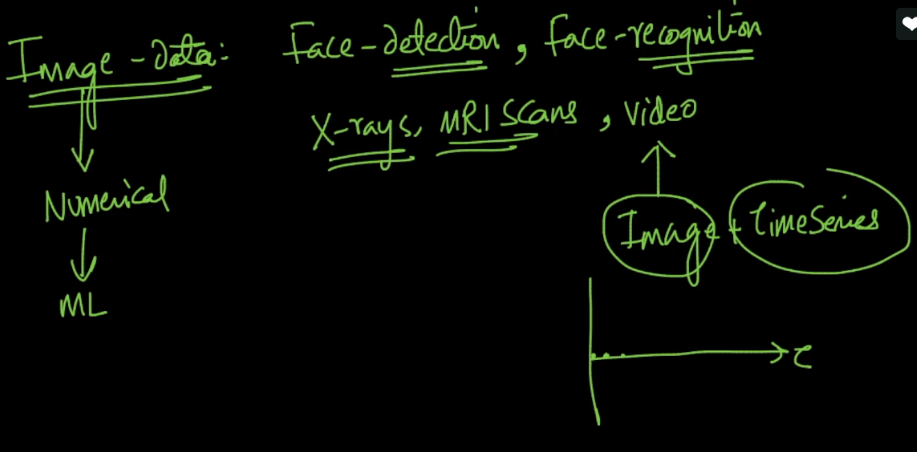
Other ex: heart rate, products sold in an hour, speech/audio.

So we’ll learn how to convert them to numerical vector so that they can be applied to any ML model.



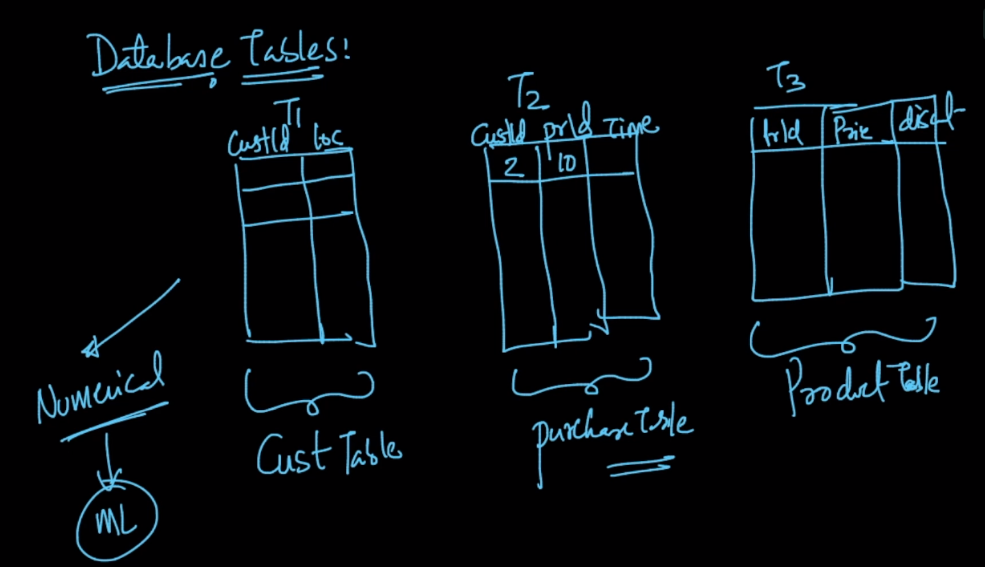
1. **Image-data:**

There are various techniques like face-detection, face recognition where images are used. And there are various types of images like X-Ray, MRI scans, video(Image + time-series). And there can be several feature eng techniques for each type of images.



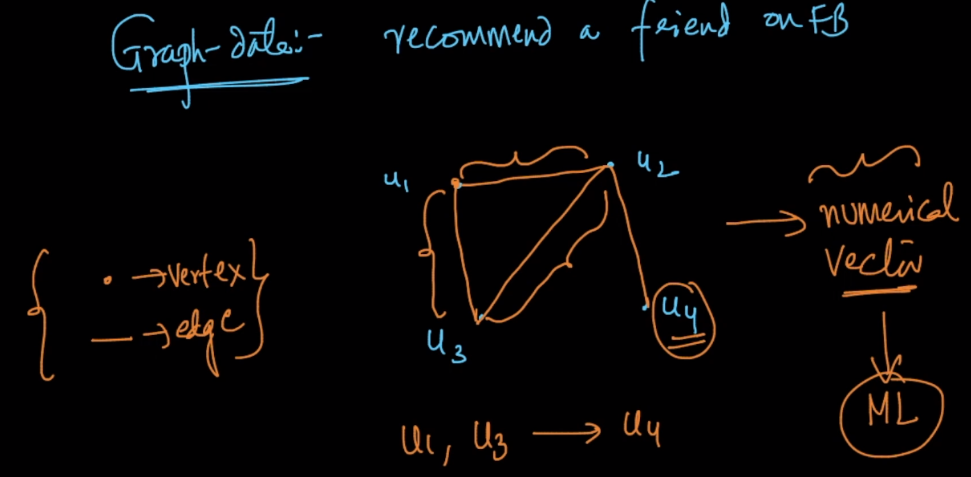
1. **Database tables:**

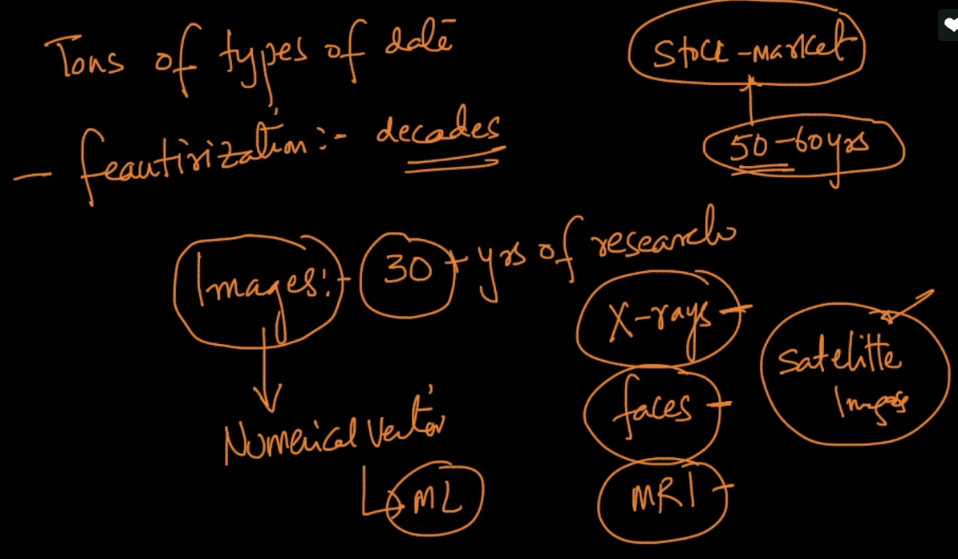
Data may be available in many table of db.

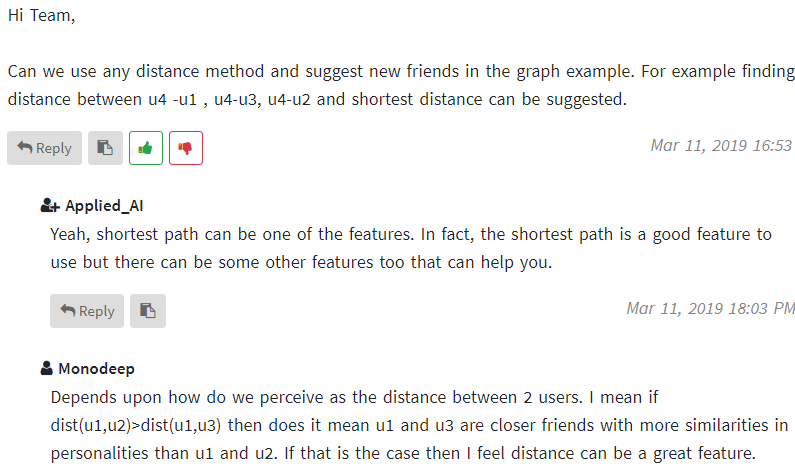


1. **Graph data:**

Like FB friend recommendataion, here each node is user, and it searches for relationship to recommend to another user. We’ll see how to convert it into numerical vector







So here we’ll cover only general purpose techniques.

