Supervised Learning

Learning from experience is the key

- Learn from past experience - In case of Machine learning it's past labeled data

Experience

----------Features---------- Response Variable / label

NO.	SIZE	COLOR	SHAPE	FRUIT NAME
1	Big	Red	Rounded shape with a depression at the top	Apple
2	Small	Red	Heart-shaped to nearly globular	Cherry
3	Big	Green	Long curving cylinder	Banana
4	Small	Green	Round to oval, Bunch shape Cylindrical	Grape
5	Medium	Red	Round	??????

Example 2

- A Credit card company receives thousands of applications for new cards. Each application contains information about an applicant,
  - Age
  - Marital status
  - Annual Income
  - Outstanding debts
  - Credit rating

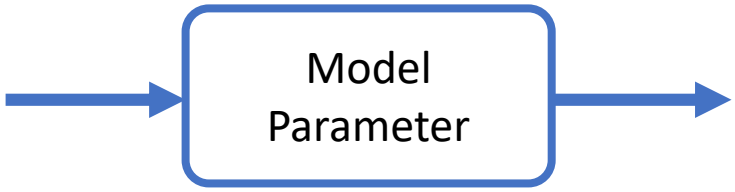
Problem: To decide whether an application should approved or rejected

Supervised Learning System

Algorithm

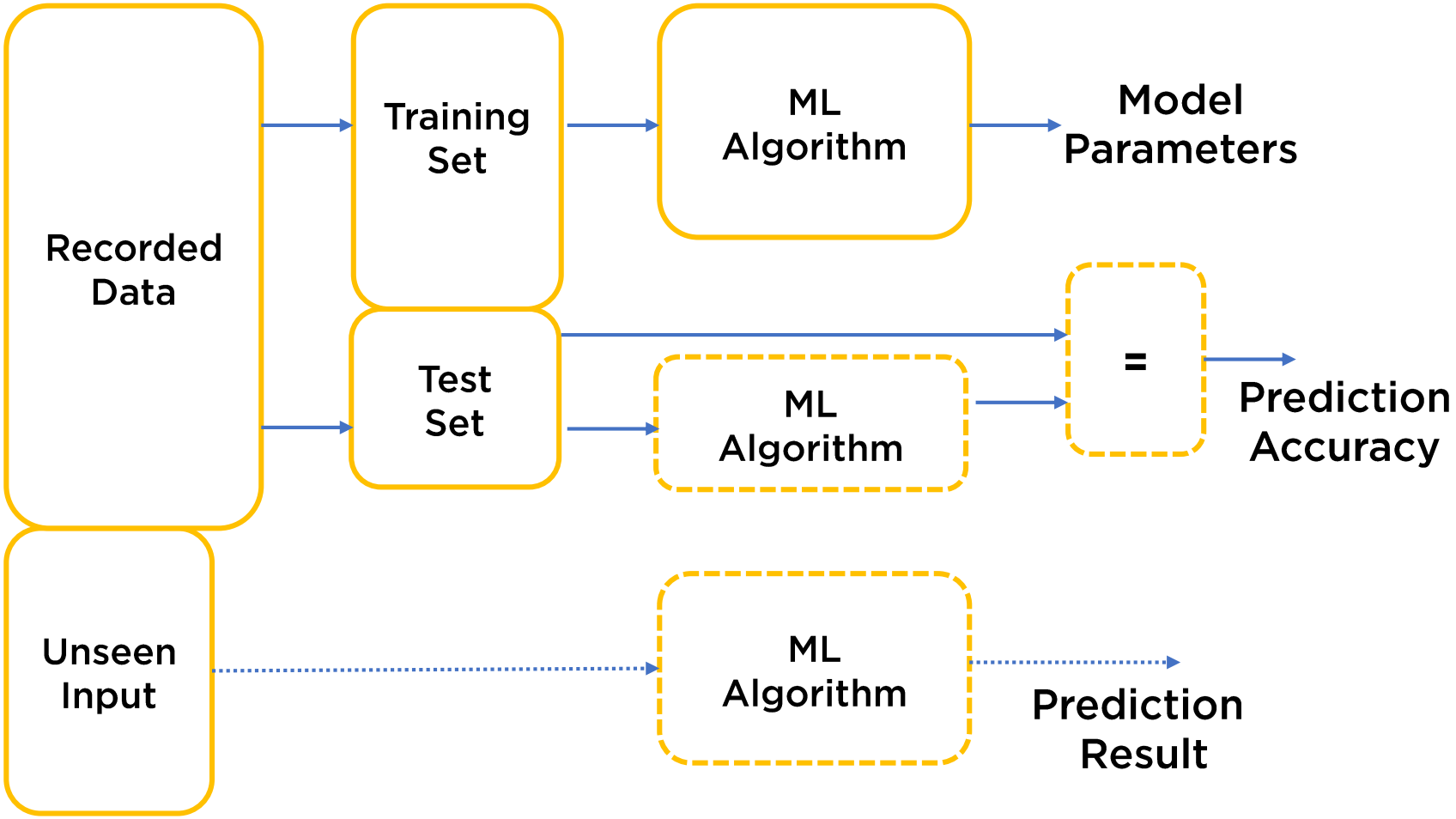
Training Phase

Production Phase



Gathering Knowledge / Creating Memories

Supervised Learning



Training and Test Dataset

- Sufficiently Variation on both data set is required for better production accuracy

- In practice, even after training, this assumption is often violated to certain degree.

Out of Syllabus Question :P

Supervised Learning

- One set of data called **Training data** consists of inputs data and correct responses corresponding to every piece of data
- Based on this training data, the algorithm has to **generalize** such that it is able to correctly (or with a low margin of error) respond to all possible inputs
- The algorithm should produce sensible outputs for inputs that weren't encountered during training.
- Also called learning from examples

Performance Measurement (P)

- **Accuracy**

$$= \frac{No\ of Correct\ Predictions}{Total\ Inputs}$$

- **Residue Sum of Squares**

$$SS = \sum_{i=\ 0}^{n} \left(y - f(x_i)\right)^2$$

- **R^2 – R Squared Value**

$$= 1 - \frac{\sum_{i=\ 0}^{n} (y - \bar{y})^2}{\sum_{i=\ 0}^{n} \left(y - f(x_i)\right)^2}$$

$$\bar{y} = Mean\ of\ sequence \qquad f(x_i) = Predicted\ Value$$

Supervised Learning

- Regression
  - Continuous variable

Example: What be price of 10 item? How much gold cost after 2 years?

- Classification
  - Discrete answers

Example: Stock trendline?, can India win world cup?

Linear Regression

Some definitions

- Dependent Variable (Single Explanatory Variable)

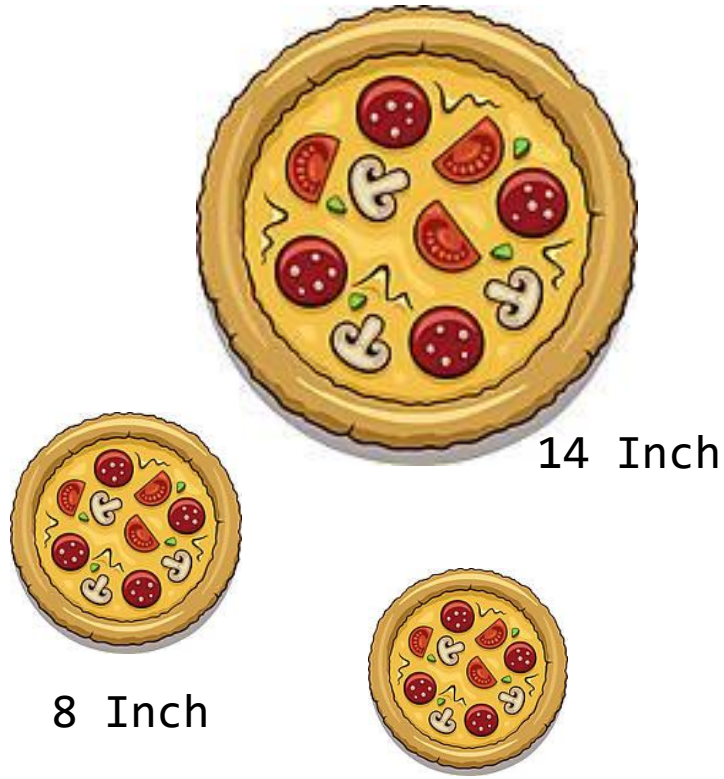
- Independent Variable (Response Variable)

Examples

- Suppose you wish to know the price of a pizza
- Estimation based on linear regression

Record	Pizza Size	Price
1	6	7
2	8	9
3	10	13
4	14	17.5
5	18	18

Menu Card



Ordinary Least Square Estimation

$$Y = mx + c$$

$$m = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

$$c = \bar{y} - m.\ \bar{x}$$

Mean Squared Error (Loss Function)

Stochastic Gradient Decent

Evaluating Performance (R Square)

$$SS_{tot} = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$SS_{res} = \sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Evaluating Performance (K-Fold Cross Validation)

```
cross_val_score

cross_val_predict
```

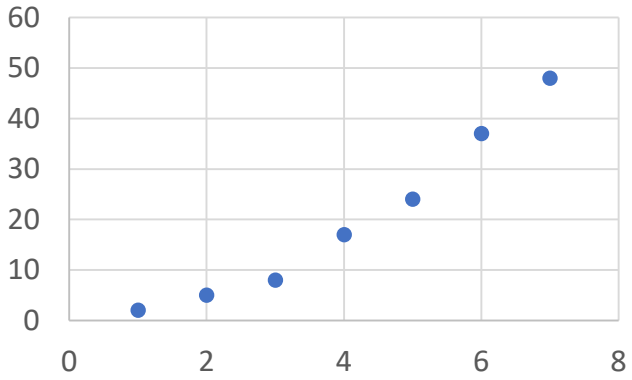
Regression.ipynb

Multivariate Linear Regression

Regression.ipynb

Non-Linear (Polynomial) Regression

Converting Non-Linear to Linear



X	Y
1	2
2	5
3	8
4	17
5	24
6	37
7	48

Overfitting of polynomial

Logistic Regression

Logistic Regression

- LOGIT Function Explanation

Performance Metrics

- Confusion Matrix
  - False Positives
  - True Negatives
- Accuracy
- Precision and Recall
- F1-Score
- ROC Curve

- Accuracy - Accuracy is the most intuitive performance measure and it is simply a ratio of correctly predicted observation to the total observations. One may think that, if we have high accuracy then our model is best. Yes, accuracy is a great measure but only when you have symmetric datasets where values of false positive and false negatives are almost same. Therefore, you have to look at other parameters to evaluate the performance of your model. For our model, we have got 0.803 which means our model is approx. 80% accurate.
- Accuracy = TP+TN/TP+FP+FN+TN

- Precision - Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. The question that this metric answer is of all passengers that labeled as survived, how many actually survived? High precision relates to the low false positive rate. We have got 0.788 precision which is pretty good.

- **Recall** (Sensitivity) - Recall is the ratio of correctly predicted positive observations to the all observations in actual class - yes. The question recall answers is: Of all the passengers that truly survived, how many did we label? We have got recall of 0.631 which is good for this model as it's above 0.5.
- Recall = TP/TP+FN

References

1. https://www.youtube.com/watch?v=IyDwQNXDWns

Ref: http://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/