

MACHINE LEARNING WORKSHOP

Sowmya S Sundaram

OVERVIEW

- Introduction to Machine Learning
- Overview of Kaggle
- Math of Linear Regression
- Linear Regression – A Code based Analysis
- Data Exploration
- Regression Competition

INTRODUCTION TO MACHINE LEARNING

- Learn from data
- Rain example
- Why is it popular now?

KAGGLE

- Your home for data science
- Find and publish data sets
- Explore and build models in a web-based data-science environment
- Work with other data scientists and machine learning engineers
- Enter competitions to solve data science challenges.

KAGGLE ACCOUNT CREATION

MACHINE LEARNING - THE CRUX

- See a bunch of data, find patterns and predict new unseen data
- More data -> more reliable prediction
- More knowledge -> better ways of capturing data
- More research -> better algorithms
- More practise -> better products :)

LINEAR REGRESSION - MATH

LINEAR REGRESSION

-

HANDS ON!

DATA EXPLORATION

DATA EXPLORATION

- Find the characteristics of data - number of features, number of classes/ range of values predicted
- Normalize data - bring everything to 0-1 or centre everything with the mean
- Missing values - fill with mean / put some constant to signify missing, remove the column in itself, search for duplicates

DATA EXPLORATION

-

HANDS ON

COMPETITION!

KEEP KODING!