# I. Intro

## 1. Welcome

1. Audience
2. Algorithms descriptions
3. Book structure
   1. 1. Background on machine learning algorithms.
   2. 2. Linear machine learning algorithms- following algorithms 1. Gradient descent. 2. Linear regression. 3. Logistic regression 4. Linear discriminant analysis
      1. 4 algorithms
         1. Gradient descent optimization procedure that may be used in the heart of many machine learning algorithms.
         2. Linear regression for predicting real values with two tutorials to make sure it really sinks in.
         3. Logistic regression for classification on problems with two categories.
         4. Linear discriminant analysis for classification on problems with more than two categories.
   3. 3. Nonlinear machine learning algorithms.
      1. Idea: make fewer assumptions about your problem and are able to learn a large variety of problem types. But this power needs to be used carefully because they can learn too well and overfit your training data.
      2. 5 algorithms
         1. Classification and regression trees the staple decision tree algorithm.
         2. Naive Bayes using probability for classification with two tutorials showing you useful ways  this technique can be used.
         3. k-Nearest Neighbors that do not require any model at all other than your dataset.
         4. Learning Vector Quantization which extends k-Nearest Neighbors by learning to compress your training dataset down in size.
         5. Support vector machines which are perhaps one of the most popular and powerful out of the box algorithms.
   4. 4. Ensemble machine learning algorithms.
      1. Idea: combine the predictions from multiple models in order to provide more accurate predictions
      2. 2 algorithms
         1. Bagging and Random Forests which are among the most powerful algorithms available.
         2. Boosting ensemble and the AdaBoost algorithm that successively corrects the predictions of weaker models.
4. What this book is NOT
5. How to best use this book
6. Summary

# II. Background

## 2. How to talk abt data in ML

1. Data as you know it
2. Statistical learning perspective
   1. Output = f(Input)
   2. OutputVariable = f(InputVariables)
   3. OutputVariable = f(InputVector)
   4. DependentVariable = f(IndependentVariables)
   5. Y =f(X)
3. Computer science perspective
   1. OutputAttribute = Program(InputAttributes)
   2. Output = Program(InputFeatures)
   3. Prediction = Program(Instance)
4. Models and Algorithms
   1. both algorithm and model can be used interchangeable
   2. model as the specific representation learned from data AND the algorithm as the process for learning it
      1. Model = Algorithm(Data)
   3. For example, a decision tree or a set of coefficients are a model
      1. and the C5.0 and Least Squares Linear Regression are algorithms to learn those respective models.
5. Summary

## 3. Algorithms Learn a Mapping from Input to Output

1. Learning a Function
   1. Y =f(X)+e
   2. This error might be error such as not having enough attributes to sufficiently characterize the best mapping from X to Y
2. Learning a function to make predictions: learn the mapping Y = f(X) to make predictions of Y for new X
3. Techniques for Learning a Function
4. Summary

## 4. Parametric and Nonparametric ML algorithms

## 5. Supervised, Unsupervised and Semi-supervised Learning

## 6. The Bias- Variance Trade-off

## 7. Overfitting and Underfitting

# III. Linear algorithms

## 8. Crash-course in spreadsheet math

## 9. Gradient Descent for ML

## 10. Linear Regression

### 11. Simple Linear Regression Tutorial

### 12. Linear Regression tutorial using Gradient Descent

## 13. Logistic Regression

### 14. Logistic Regression Tutorial

## 15. Linear Discriminant Analysis

### 16. Linear Discriminant Analysis Tutorial

# IV. Nonlinear algorithms

# V. Ensemble algorithms