# ML Programming Lab Tutorial Outline

**Tutorial 1**

* + Letter counts and frequencies
  + Fitting lines to continuous data.
  + Linear regression from scratch
  + 3D plotting with generated data.
  + Explore and visualize airport data.

**Tutorial 2**

* + Data handling and exploration. Data types.
  + Ordinal data.
  + Data encoding.
  + Binary encoding.
  + Pros cons of binary encoding.

**Tutorial 3**

* + Booth function gradient descent.
  + Linear regression using Normal Equations.
    - Gaussian elimination.
    - QR decomposition.
  + Numpy SLE
  + Time Series exploration
    - Train/test split, scaling data.

**Tutorial 4**

* + Part 1. Optimization and loss classes for MSE and cross-entropy.
  + Part 2. Linear regression class with Fit and Predict
    - Using the Optimization and Loss.
    - Comparison against the Sklearn linear regression.
  + Part 3. Logistic regression class with Fit and Predict functions.
    - Using Entropy and Newton’s Method.
    - Precision, recall, F-score, confusion matrix.
  + Part 4. Discriminant Analysis.
    - 2 classes LDA and QDA with their associated Fit and Predict functions.
    - Comparison with Sklearn implementation.

**Tutorial 5**

* + Part 1. Gradient Descent for Ridge Regression.
    - MSE.
      * L1 w/MSE.
      * L2 w/MSE.
    - Cross Entropy.
      * L1.
      * L2.
    - Weighted Cross Entropy.
    - ElasticNet.
  + Part 2: Regularized Logistic Regression.

**Tutorial 6**

* + Part 1. Variable Selection.
    - Forward Search.
    - Backward Search.
  + Part 2. Regularization and Hyperparameter Search.
    - Ridge Regression
    - Lasso
    - ElasticNet
    - Grid Search
    - Random Search

**Tutorial 7**

* + Part 1. Distance Metrics.
    - Levenshtein Distance.
    - Minkowski Metric.
      * Manhattan.
      * Euclidean.
      * Chebyshev.
  + Part 2. K-Nearest Neighbor Classifier.
    - Standard Scaling.
    - Hyperparameter search.
    - Precision, recall, F1-score.
    - Plot decision boundaries.
  + Part 3. Over/under fitting.
    - Polynomial Regression.
    - Polynomial feature transformer.

**Tutorial 8**

* + Part 1. Bag of Words and TF-IDF representation.
    - Tokenization of text.
    - Removing stop words and tags.
    - Naive Bayes.
  + Part 2. N-gram Language Model.

**Tutorial 9**

* + Part 1. Perceptron. Visualization of different decision boundaries.
  + Part 2. Neural Network.
    - Feed forward.
    - Back propagation.
    - Stochastic Gradient Descent.
  + Part 3. Multi Layer Perceptron Classifier.
    - MLP Classification.
      * K-fold cross-validation.
      * Random search.
    - MLP Regressor.

**Tutorial 10**

* + Part 1. Decision Trees.
    - Appropriate stopping criterion.
    - Grid or random search with Sklearn.
  + Part 2. NLP.
    - Continuous bag of words
    - Neural Network with PyTorch.

**Tutorial 11**

* + Part 1. Support Vector Machine.
    - Pegasos Algorithm
    - Mini batch Pegasos
    - Dual Coordinate Descent.

**Tutorial 12**

* + Part 1. K-means.
    - Evaluation methods:
      * Elbow.
      * Average silhouette.
      * Gap statistic.
    - Principal Component Analysis PCA.
  + Part 2. Gaussian Mixtures.