Pattern Classification & Machine Learning - Project 1
Finding the Higgs Boson

Apolline Lafarge, Loic Ottet, Dimitrios Sarigiannis
Department of Computer Science, EPFL Lausanne, Switzerland

Abstract—**In this report, we summarize our findings for the first project of the machine learn- ing course taught at the EPFL. The goal of the project is the computation as well as the error estimation of predictions of a classification dataset of the CERN Higgs boson. Key aspects of the project are exploratory data analysis, feature processing, the implementation of several regression and classification machine learning methods and the comparison of different models using cross-validation.**

I. Introduction

text

II. Data Analysis

A. Data Description

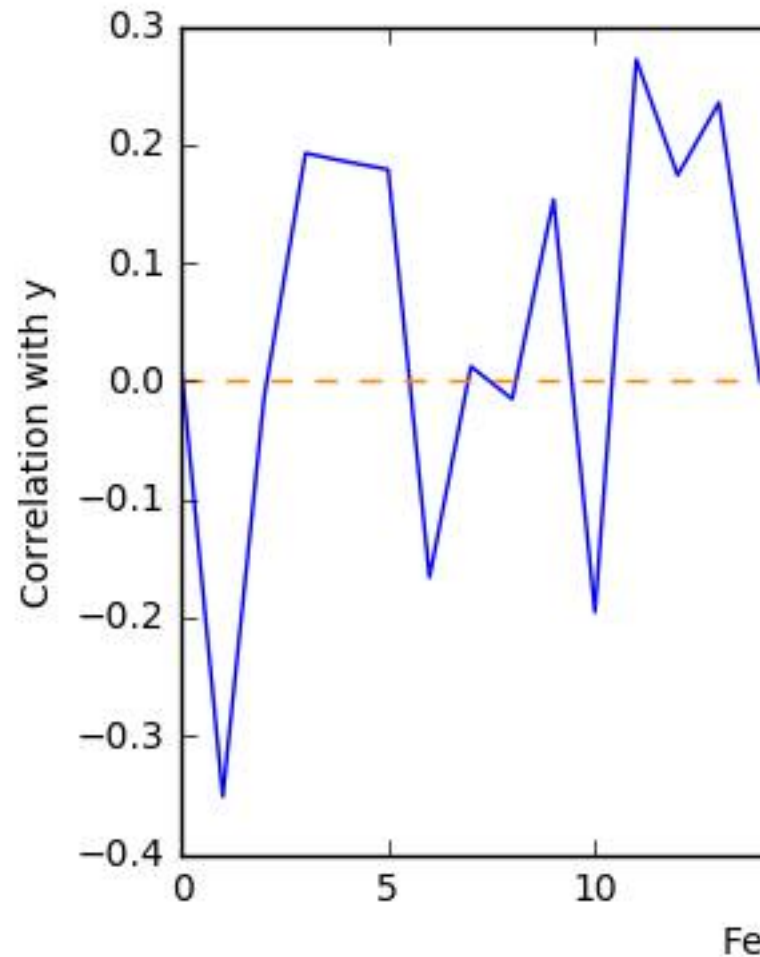text text text text text text text text text

B. Data Analysis

text text text text text text text text text text

C. Rank Deficiency

text text text text text text text text text text text text

D. Feature processing

text text text text text text text text text text

III. ML Regression Methods

For the regression task, we have implemented 3 methods: least-squares using gradient descent, least-squares using ridge regression, least-squares using normal equations and ridge regression using normal equations

A. Least Squares using Gradient Descent

text text text text

B. Least Squares using Stochastic Gradient Descent

text text text text

C. Least Squares using normal equations

Given that the input matrix X is not ill-conditioned, we are able to apply least-squares using normal equations.

D. Ridge Regression using normal equations

text text text text

IV. ML Classification Methods

A. Logistic Regression using Gradient Descent

text text text text text

B. Regularized Logistic Regression using Gradient Descent

text text text text

V. Performance Analysis

text

VI. Summary

text