

Compiler Provenance

Homework 1 - Machine Learning
Engineering in Computer Science
"La Sapienza" University of Rome

Costa Marco 1691388

6 November 2019

Contents

1	Introduction	3
2	Dataset	3
3	Find best classifier	3
4	Predictions	3
5	Conclusions	3

1 Introduction

The purpose of this homework is to apply machine learning algorithms to solve the compiler provenance problem. It consists of identifying the **compiler** and the **optimization** who produced a given binary code of a function. So we have two types of classification problems:

- **Binary classification:** optimization can be *HIGH (H)* or *LOW (L)*
- **Multi-class classification:** compiler can be *gcc*, *icc* or *clang*

2 Dataset

3 Find best classifier

4 Predictions

5 Conclusions