

# Relatório TP1 SBS-MLFA

Luis Freitas; Daniel Pereira

November 10, 2020

**Abstract**

## **1 Introdução**

## **2 Metodologias**

### **2.1 CRISP-DM**

### **2.2 Modelo Baseados em Árvores**

### **3   Arquiteturas e Ferramentas**

## 4 1º Dataset

## **4.1 Contexto do Dataset**

O dataset utilizado neste projeto é referente ao nível de incidentes rodoviários na zona de Braga e contém algumas features que representam a magnitude do atraso que se verifica a cada hora, a temperatura, a pressão atmosférica, a velocidade do vento e ainda outras features.

O principal objetivo é aprimorar um modelo de Machine Learning Baseado em Árvores mas também fazer uma análise exploratoria completa dos dados para se retirar algumas informações importantes do dataset. Outro ponto passa por conseguir preparar o dataset com técnicas de engenharia de dados para se obter os melhores resultados possíveis. Este dataset é referente a uma competição da plataforma Kaggle

Em relação as features do dataset estas são apresentadas da seguinte forma:

1)

## **4.2 Análise e Compreensão dos Dados**

## **4.3 Modelação**

## **4.4 Análise de Resultados**

# **5 Dataset de Competição**

- 5.1 Contexto do Dataset
- 5.2 Análise e Compreensão dos Dados
- 5.3 Modelação
- 5.4 Análise de Resultados
- 6 Conclusão