

# Data Mining & Machine Learning F20DL

Group 4

Lewis Wilson, Sam Fay-Hunt, Kamil Szymczak, Chun Man

November 30, 2020

# Contents

<b>1</b>	<b>Variation in performance with size of the training and testing sets</b>	<b>1</b>
<b>2</b>	<b>Variation in performance with change in the learning paradigm (Decision Trees versus Neural Nets)</b>	<b>2</b>
<b>3</b>	<b>Variation in performance with varying learning parameters in Decision Trees</b>	<b>3</b>
3.1	J48 . . . . .	3
3.2	Random Forest . . . . .	4
<b>4</b>	<b>Variation in performance with varying learning parameters in Neural Networks</b>	<b>5</b>
4.1	Linear Classifier . . . . .	5
4.2	Multilayer Perceptron . . . . .	6
<b>5</b>	<b>Variation in performance according to different metrics (TP Rate, FP Rate, Precision, Recall, F Measure, ROC Area)</b>	<b>7</b>
<b>Appendices</b>		<b>8</b>
<b>A</b>	<b>Appendix A</b>	<b>8</b>

# 1 Variation in performance with size of the training and testing sets

## 2 Variation in performance with change in the learning paradigm (Decision Trees versus Neural Nets)

### 3 Variation in performance with varying learning parameters in Decision Trees

#### 3.1 J48

## 3.2 Random Forest

## 4 Variation in performance with varying learning parameters in Neural Networks

### 4.1 Linear Classifier

## 4.2 Multilayer Perceptron



## 5 Variation in performance according to different metrics (TP Rate, FP Rate, Precision, Recall, F Measure, ROC Area)

# Appendices

## A Appendix A

Figure 1: Accuracy vs Created - Random Forest

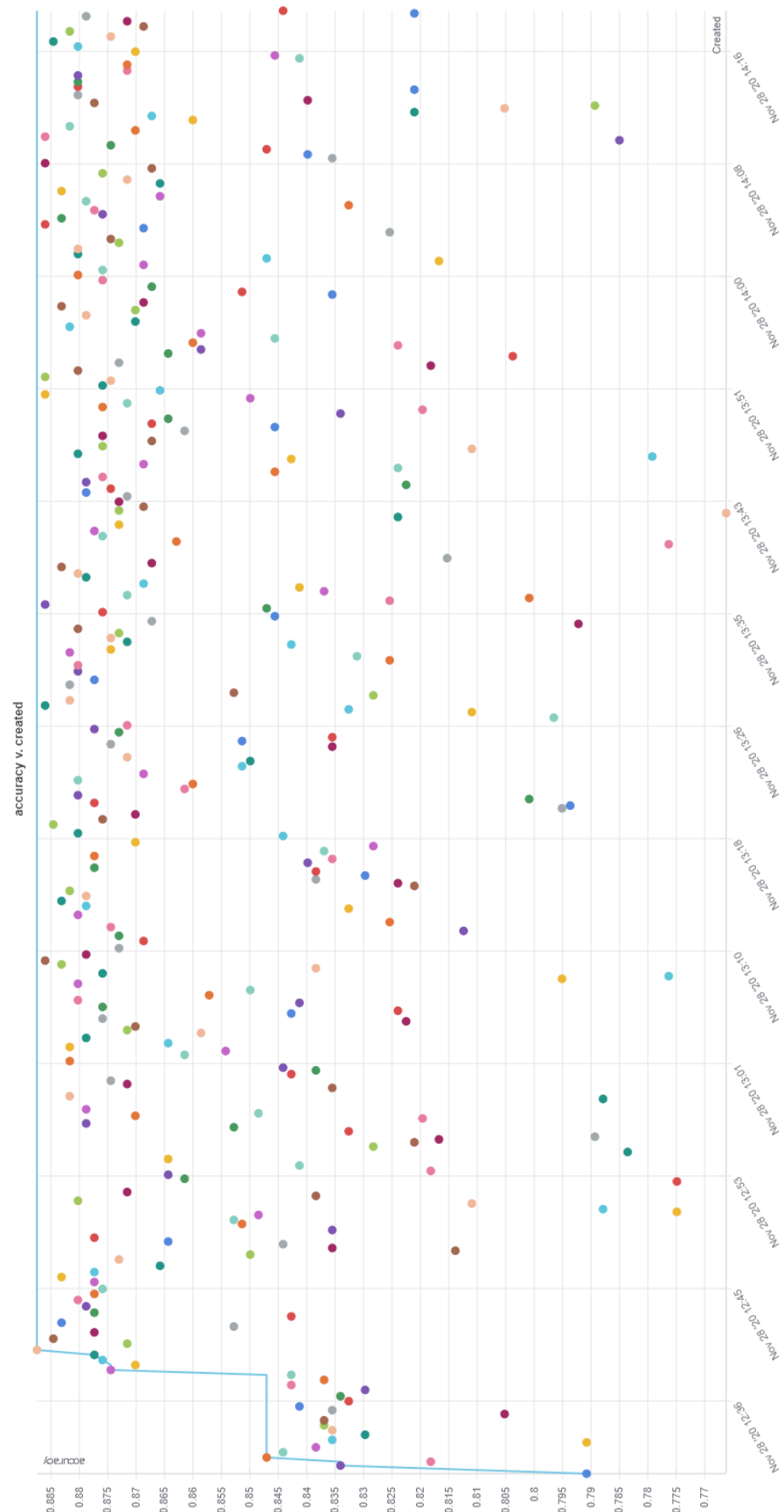


Figure 2: Parameter Importance - Random Forest

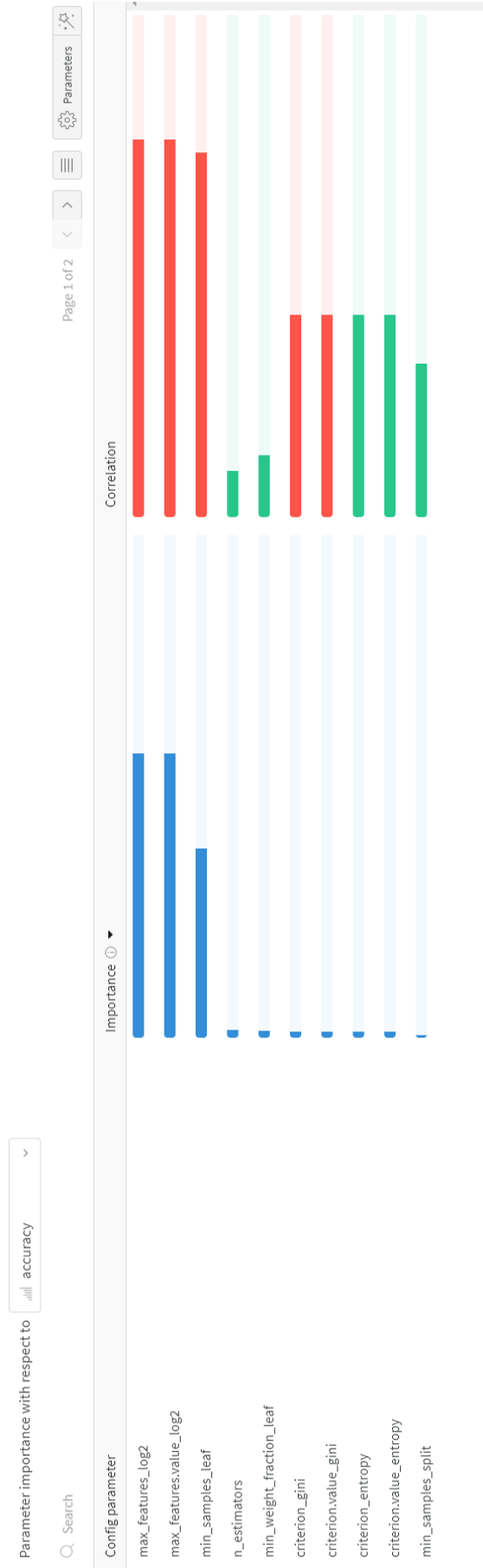


Figure 3: Random Forest Parameters

