

UNIVERSIDAD DE ALCALÁ

Escuela Politécnica Superior

**GRADO EN INGENIERÍA INFORMÁTICA**

Trabajo Fin de Grado

**Open Source Data Mining Tools for Association Rules and  
Subgroup Discovery**

**Autor:** Javier López Ruiz

Director: Daniel Rodríguez García

**TRIBUNAL:**

**Presidente:**

**Vocal 1º:**

**Vocal 2º:**

**CALIFICACIÓN:** .....

**FECHA:** .....





UNIVERSITY OF ALCALÁ

DEPARTMENT OF COMPUTER SCIENCE

# Open Source Data Mining Tools for Association Rules and Subgroup Discovery

Dissertation conducted by

**Javier López Ruiz**

Dissertation supervised by

**Daniel Rodríguez García**

September 2014



# Resumen

Texto

**Palabras clave:** palabra



# Abstract

Texto

**Keywords:** word





# *Acknowledgements*

The acknowledgements and the people to thank go here, don't forget to include your project advisor. . .



*“If you torture the data long enough, it will confess.”*

Ronald Coase



# Contents

<b>Resumen</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>List of Figures</b>	<b>xiii</b>
<b>List of Tables</b>	<b>xv</b>
<b>Abbreviations</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Objectives . . . . .	1
1.2 Methodology . . . . .	1
1.3 Resources . . . . .	1
1.4 Work structure . . . . .	1
<b>2 Weka</b>	<b>3</b>
2.1 Overview . . . . .	3
2.2 Features . . . . .	3
2.3 Demo . . . . .	3
2.4 Final thoughts . . . . .	3
<b>3 R Environment</b>	<b>5</b>
3.1 Overview . . . . .	5
3.2 Features . . . . .	5
3.3 Demo . . . . .	5
3.4 Final thoughts . . . . .	5
<b>4 Rapidminer</b>	<b>7</b>
4.1 Overview . . . . .	7
4.2 Features . . . . .	7
4.3 Demo . . . . .	7
4.4 Final thoughts . . . . .	7
<b>5 KNIME</b>	<b>9</b>
5.1 Overview . . . . .	9

---

5.2	Features . . . . .	9
5.3	Demo . . . . .	9
5.4	Final thoughts . . . . .	9
<b>6</b>	<b>JCLEC</b>	<b>11</b>
6.1	Overview . . . . .	11
6.2	Features . . . . .	11
6.3	Demo . . . . .	11
6.4	Final thoughts . . . . .	11
<b>7</b>	<b>KEEL</b>	<b>13</b>
7.1	Overview . . . . .	13
7.2	Features . . . . .	13
7.3	Demo . . . . .	13
7.4	Final thoughts . . . . .	13
<b>8</b>	<b>Cortana</b>	<b>15</b>
8.1	Overview . . . . .	15
8.2	Features . . . . .	15
8.3	Demo . . . . .	15
8.4	Final thoughts . . . . .	15
<b>9</b>	<b>Orange</b>	<b>17</b>
9.1	Overview . . . . .	17
9.2	Features . . . . .	17
9.3	Demo . . . . .	17
9.4	Final thoughts . . . . .	17
<b>10</b>	<b>Conclusions and feature work</b>	<b>19</b>
10.1	Conclusions . . . . .	19
10.2	Future work . . . . .	19
<b>A</b>	<b>Learning resources</b>	<b>21</b>
	<b>Bibliography</b>	<b>23</b>

# List of Figures





# List of Tables



# Abbreviations

**LAH** List Abbreviations **Here**



# Chapter 1

## Introduction

1.1 Objectives

1.2 Methodology

1.3 Resources

1.4 Work structure



## Chapter 2

# Weka

### 2.1 Overview

### 2.2 Features

### 2.3 Demo

### 2.4 Final thoughts





## Chapter 3

# R Environment

### 3.1 Overview

### 3.2 Features

### 3.3 Demo

### 3.4 Final thoughts



## Chapter 4

# Rapidminer

### 4.1 Overview

### 4.2 Features

### 4.3 Demo

### 4.4 Final thoughts



## Chapter 5

# KNIME

### 5.1 Overview

### 5.2 Features

### 5.3 Demo

### 5.4 Final thoughts



## Chapter 6

# JCLEC

### 6.1 Overview

### 6.2 Features

### 6.3 Demo

### 6.4 Final thoughts





## Chapter 7

# KEEL

### 7.1 Overview

### 7.2 Features

### 7.3 Demo

### 7.4 Final thoughts



## Chapter 8

# Cortana

### 8.1 Overview

### 8.2 Features

### 8.3 Demo

### 8.4 Final thoughts



## Chapter 9

# Orange

### 9.1 Overview

### 9.2 Features

### 9.3 Demo

### 9.4 Final thoughts



## Chapter 10

# Conclusions and feature work

### 10.1 Conclusions

### 10.2 Future work





## Appendix A

### Learning resources



# Bibliography