

UNIVERSIDAD DE GRANADA
E.T.S. de Ingenierías Informática y de Telecomunicación



Departamento de Ciencias de la
Computación e Inteligencia Artificial

Inteligencia de Negocio



PRÁCTICA 4

Escalabilidad en Grandes Conjuntos de Datos

Christian Andrades Molina

ÍNDICE



- 1. Especificaciones del hardware**
- 2. Estudio de escalabilidad**
- 3. Estratificación**

1. Especificaciones del hardware

Tipo de computadora : Equipo basado en x64 ACPI

Procesador : Intel(R) Core(TM) i7-5500U CPU @ 2.40GHz

Memoria RAM : 4GB

Sistema operativo : Windows 10 Home Edition, 64 bits

Placa de video : Intel(R) HD Graphics 5500

El comando utilizado para Protein es:

```
java -Xmx2048M weka.classifiers.trees.J48 -t  
resultados_20_protein.arff -T datos_test_protein.arff -i >  
modelo_resultado_20_protein.doc
```

```
java weka.classifiers.trees.J48 -t data/kddcup994.arff -T  
data/kddcup99_prueba.arff > resultado_4_kddcup99.doc
```

2. Estudio de escalabilidad

3. Estratificación