



UNIVERSIDAD DE GRANADA

APRENDIZAJE AUTOMÁTICO
GRADO EN INGENIERÍA INFORMÁTICA

PRÁCTICA 3

PROGRAMACIÓN

Autor

Vladislav Nikolov Vasilev

Rama

Computación y Sistemas Inteligentes



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS INFORMÁTICA Y DE
TELECOMUNICACIÓN

CURSO 2018-2019

Índice

1. PROBLEMA DE REGRESIÓN	2
1.1. Descripción del problema	2
1.2. Análisis de los datos	2
2. PROBLEMA DE CLASIFICACIÓN	4
Referencias	5

1. PROBLEMA DE REGRESIÓN

1.1. Descripción del problema

En este problema vamos a trabajar con el conjunto de datos *Airfoil Self-Noise*, el cuál ha sido proporcionado por la NASA, y contiene los resultados de haber realizado un conjunto de pruebas aerodinámicas y acústicas en un túnel de viento sobre perfiles alares de dos y tres dimensiones.

El conjunto de datos está compuesto por 1503 filas y 6 columnas, los valores de las cuáles son todos números reales. Los datos de las 5 primeras columnas se corresponden con los datos de entrada, y la última columna se corresponde con la información de salida. A continuación se puede ver que representa cada uno de los atributos de forma ordenada:

1. Frecuencia, medida en Hz .
2. Ángulo de ataque (ángulo que forman la cuerda geométrica de un perfil alar con la dirección del aire incidente), medida en grados.
3. Longitud de la cuerda del perfil alar, medida en metros.
4. Velocidad *free-stream*, medida en metros por segundo.
5. Distancia de desplazamiento de succión, medida en metros.
6. Nivel de presión sonora, medida en dB .

1.2. Análisis de los datos

Antes de comenzar con todo el proceso de elección y selección de un modelo lineal, vamos a pararnos un momento para analizar los datos de los que disponemos con el fin de obtener más información sobre el problema.

Lo primero que tenemos que hacer es cargar los datos. Para ello, vamos a usar una función genérica que nos permita leer ficheros de datos y obtener un *DataFrame* que podamos usar luego. Vamos a ver como sería esta función:

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```
# Establecer la semilla que vamos a utilizar
np.random.seed(1)

def read_data_values(in_file, separator=None):
    """
    Funcion para leer los datos de un archivo

    :param in_file Archivo de entrada
    :param separator Separador que se utiliza en el archivo
        (por defecto None)

    :return Devuelve los datos leidos del archivo en un DataFrame
    """

    # Cargar los datos en un DataFrame
    # Se indica que la primera columna no es el header
    if separator == None:
        df = pd.read_csv(in_file, header=None)
    else:
        df = pd.read_csv(in_file, sep=separator, header=None)

    return df
```

Con la función ya mostrada, vamos a cargar los datos y mostrar los primeros valores de la muestra, para tener una idea de como serán los datos:

Con la función ya mostrada, vamos a cargar los datos y mostrar los primeros valores de la muestra, para tener una idea de como serán los datos:

```
In [2]: df = read_data_values('datos/airfoil_self_noise.dat', separator='\t')

# Asignamos nombres a las columnas (según los atributos)
column_names = ['Frequency', 'Angle of attack', 'Chord length',
                'Free-stream velocity', 'SSD thickness',
                'Sound Pressure']
df.columns = column_names

# Mostrar primeros valores de los datos
df.head()
```

2. PROBLEMA DE CLASIFICACIÓN

Referencias

- [1] Texto referencia
<https://url.referencia.com>