



Universitat de Lleida

Exercicis d'ARA

Bayesian Inference

Realitzat per:

Oriol Alàs Cercós, Marta Albets Mitjaneta, Sergi Simón Balcells

Data de lliurament:

8 de Juny de 2020

Universitat de Lleida
Escola Politècnica Superior
Grau en Enginyeria Informàtica
Aprenentatge i Raonament Automàtic

Professorat:
Ramón Béjar

Índex

1	Introducció	2
2	Parseig de dades	2
2.1	Partició de les dades	2
3	Models	2
A	Gràfics	3

Índex de taules

1	Mètodes utilitzats segons cada atribut	2
---	--------------------------------------------------	---

Índex de figures

1	(a) Scatter de l'atribut price continu (b) Scatter de l'atribut price discret	3
2	(a) Scatter de l'atribut reviews continu (b) Scatter de l'atribut reviews discret	3
3	(a) Scatter de l'atribut latitude continu (b) Scatter de l'atribut latitude discret	3
4	(a) Scatter de l'atribut longitude continu (b) Scatter de l'atribut longitude discret	4

1 Introducció

2 Parseig de dades

2.1 Partició de les dades

Per tal de passar els atributs que tenen valors reals a valors discrets, hem utilitzat la llibreria `pandas`, que ens permet llegir el csv d'una manera molt còmode i partir-lo de diferents maneres. Entre aquestes maneres, hem vist que dos mètodes eren bastant bons:

- **cut**. Divideix el dataset en diferents intervals utilitzant el rang de valors.
- **qcut**. Divideix el dataset mitjançant quantiles, és a dir, tots els intervals tindran el mateix nombre de dades.

Per tal d'arribar a la conclusió de quin mètode era el millor per tal de categoritzar els atributs, s'ha utilitzat la llibreria `matplotlib` per tal de visualitzar les dades de manera gràfica. S'ha realitzat un gràfic ¹per cada atribut a categoritzar sobre l'*overall_satisfaction* i després categoritzant l'atribut. Aquest procediment s'ha realitzar bastants cops fins trobar una partició que fos representativa.

Atribut	Mètode	Num. Divisions
price	cut	15
reviews	cut	10
latitude	cut	20
longitude	cut	20

Taula 1: Mètodes de partició utilitzats segons cada atribut

3 Models

¹Es poden trobar a la secció A de l'apèndix.

A Gràfics

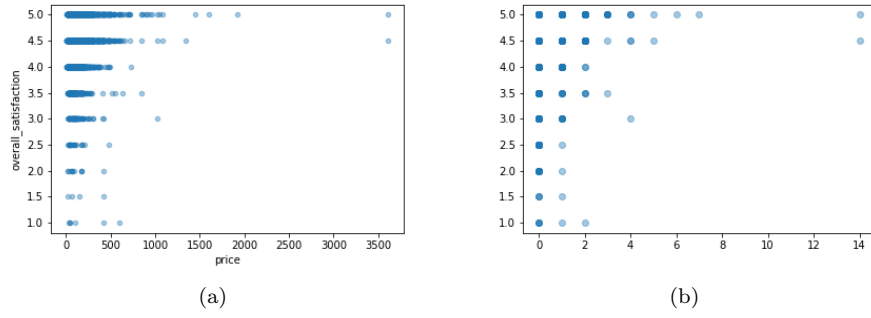


Figura 1: (a) Scatter de l'atribut price continu (b) Scatter de l'atribut price discret

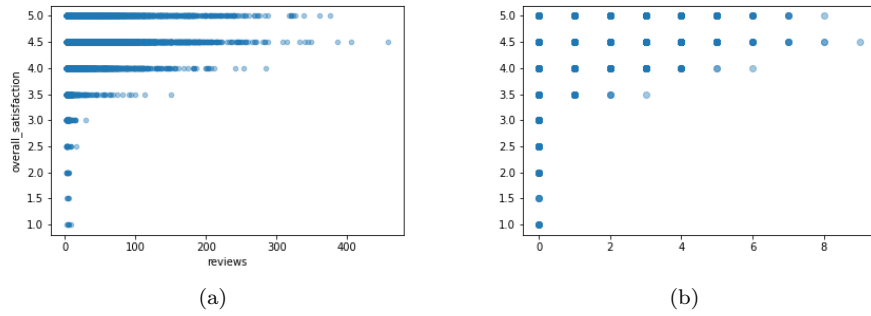


Figura 2: (a) Scatter de l'atribut reviews continu (b) Scatter de l'atribut reviews discret

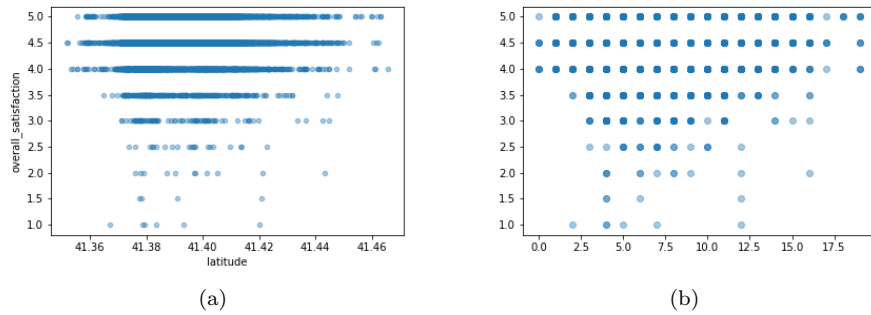
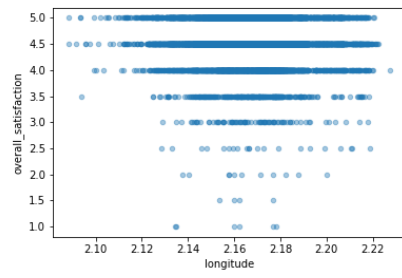
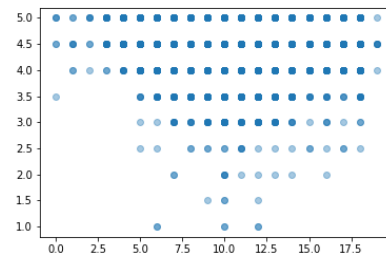


Figura 3: (a) Scatter de l'atribut latitude continu (b) Scatter de l'atribut latitude discret



(a)



(b)

Figura 4: (a) Scatter de l'attribut longitude continu (b) Scatter de l'attribut longitude discret