**VŠB – Technická univerzita Ostrava Fakulta elektrotechniky a informatiky Katedra informatiky**

**Metody analýzy dat**

**Ostrava 2020 Vojtěch Prokop**

**VŠB – Technická univerzita Ostrava Fakulta elektrotechniky a informatiky Katedra informatiky**

**Analýza marketingové kampaně Portugalské bankovní instituce**

Obsah

[1 Úvod 4](#_Toc59909284)

[2 Informace o datové sadě 5](#_Toc59909285)

[2.1 Popis atributů 5](#_Toc59909286)

[2.2 Datové typy atributů 6](#_Toc59909287)

[3 detailní Rozbor atributů 7](#_Toc59909288)

[4 shlukování 8](#_Toc59909289)

[5 klasifikace 9](#_Toc59909290)

[výsledky analýzy 10](#_Toc59909291)

[Reference 11](#_Toc59909292)

# Úvod

Datová sada reprezentuje informace nashromážděné v průběhu přímé marketingové kampaně, kterou prováděla portugalská bankovní instituce. Marketingová kampaň byla založena na telefonátech mezi bankou a jejími klienty.

# Informace o datové sadě

Vstupní datová sada obsahuje **41 118 záznamů**. Každý záznam z datové sady obsahuje 21 sloupců. Zároveň atributy datové sady jsou jak numerické, tak kategoriální, kdy mezi kategoriemi najdeme označení ***unknown*** reprezentující chybějící hodnoty.

## Popis atributů

* **Age** – věk klienta.
* **Job** – typ zaměstnání klienta.
* **Marital** – rodinný stav klienta.
* **Education** – vzdělání klienta.
* **Default** – informace o kreditní kartě klienta.
* **Housing** – informace popisující, zda klient má úvěr na byt.
* **Loan** – informace popisující, zda klient má osobní půjčku.
* **Contact** – typ komunikace s klientem.
* **Month** – měsíc, kdy byl naposled klient kontaktován.
* **Day\_of\_week –** den v týdny, kdy byl klient naposled kontaktován.
* **Duration** – doba trvání poslední komunikace s klientem.
* **Campaign** – informace popisující kolikrát byl klient kontaktován v průběhu kampaně.
* **Pdays** – kdy naposled byl klient kontaktován v předešlé kampani.
* **Previous** – počet kontaktování klienta v předešlé kampani.
* **Poutcome** – výsledek předešlé kampaně.
* **Emp.var.rate** – změna zaměstnanosti v čtvrtletí.
* **Cons.price.idx** – index spotřebitelských cen (měsíční indikace).
* **Cons.conf.idx** – index spotřebitelské důvěry (měsíční indikace).
* **Euribor3m** – evropská referenční úroková sazba (denní indikace).
* **Nr.empoyed** – počet zaměstnaných lidí (čtvrtletní indikace)
* **Y** – výstupní atribut reprezentující, zda klient provedl termínovaný vklad.

## Datové typy atributů

### Rozdělení atributů do skupin

Všechny atributy jsou rozděleny do 5 skupin:

* Atributy reprezentující informace o klientovi
* Atributy reprezentující informace o posledním kontaktu s klientem
* Atributy reprezentující další informace
* Atributy reprezentující sociální a ekonomické hodnoty
* Výstupní atribut

### Typy atributů:

|  |  |
| --- | --- |
| Typ | Četnost |
| Kategoriální | 10 |
| Numerické | 10 |
| Binární | 1 |

# detailní Rozbor atributů

## Kategoriální atributy

### Job

## Age

## Job

## Marital

## Education

## Default

## Housing

## Loan

## Contact

## Month

## Day\_of\_week

## Duration

## Campaign

## Pdays

## Previous

## Poutcome

## Emp.var.rate

## Cons.price.idx

## Cons.conf.idx

## Euribor3m

## Nr.empoyed

## Y

# shlukování

# klasifikace

# výsledky analýzy

# Reference

1. *Algoritmus CYK* [online]. 2020 [cit. 2020-12-01]. Dostupné z https://cs.qaz.wiki/wiki/CYK\_algorithm
2. Bezkontextová gramatika. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2020 [cit. 2020-12-01]. Dostupné z: https://cs.wikipedia.org/wiki/Bezkontextov%C3%A1\_gramatika
3. Chomského normální forma. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2020 [cit. 2020-12-01]. Dostupné z: https://cs.wikipedia.org/wiki/Chomsk%C3%A9ho\_norm%C3%A1ln%C3%AD\_forma
4. JANČAR, Doc. RNDr. Petr. *Teoretická informatika: učební text* [online]. Ostrava 2007: Ediční středisko VŠB-TUO, 2007 [cit. 2020-12-01]. ISBN 978-80-248-1487-2. Dostupné z: http://www.cs.vsb.cz/sawa/ti/materialy/ti.pdf
5. *Bezkontextové gramatiky* [online]. [cit. 2020-12-01]. Dostupné z: http://www.cs.vsb.cz/kot/soubory\_animaci/a-cfg\_derivace.pdf