IFC-Bank of Italy Workshop on "Data Science in Central Banking: Applications and tools"

14-17 February 2022

Mailbot – optimising the process of answering statistical queries[1]

Daphne Aurouet, Nina Blatnik, Samo Boh, Andrea Colombo, Almir Delic,
Jordi Gutiérrez, Gavril Petrov and Kristine Rikova,
European Central Bank

[1] This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the Bank of Italy, the BIS, the IFC or the other central banks and institutions represented at the event.

EUROPEAN CENTRAL BANK

EUROSYSTEM

MailBot

Optimizing the process of answering statistical queries

15/02/2022, IFS workshop

Daphne Aurouet, Nina Blatnik, Samo Boh, **Andrea Colombo**,

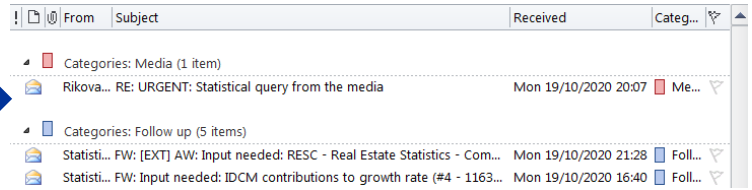
Almir Delic, Jordi Gutiérrez, **Gavril Petrov**, Kristine Rikova

Current process for handling statistical queries

EUROPEAN CENTRAL BANK
EUROSYSTEM

Statistical Information Request

Manually assign responsible team



Aspects to improve

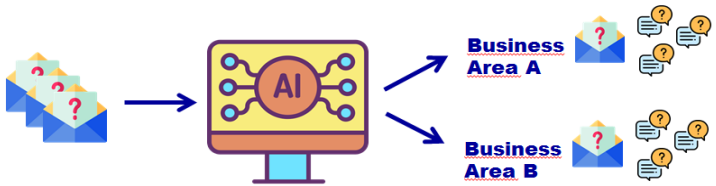
1. Slow identification; sometimes wrong BA identified

2. Similar queries answered multiple times

3. Manual handling of internal procedures

Solution

1. Classification of BAs
2. Identifying similar queries
3. Automatisation of procedures



Picture sources:
https://www.ranzey.com/
https://www.synthesizerguide.com/synthesizers/how-to-choose-a-synthesizer/
https://www.flaticon.com

Knowledge database

Simple structure ⟶ **Question** ✚ **Answer** ✚ **Responsible BA**

Text pre-processing and feature extraction

Data preparation

- Assigning IDs
- Parsing of emails – picking up original question and final reply
- Remove confidential information – greetings, signatures
- General string cleaning

NLP preprocessing

- Strip punctuation
- Tokenize
- Remove stopwords / other email vocabulary
- Lemmatizing
- Keep only nouns, pronouns, verb

Classification of BAs

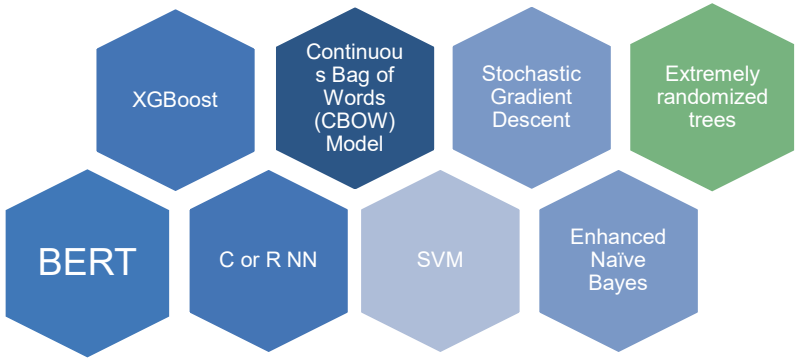
Subject: Average exchange rates

I would like to know the annual average exchange rate for the year 2019, for Canadian dollar to Euros and American dollar to Euros.

92% **ESSA Exchange rates statistics**

4% **Monetary statistics**

1% **Market Data Provision**



Extremely Randomised trees + **HEURISTICS**

Similarity of queries

Subject: Average exchange rates

I would like to know the annual average exchange rate for the year 2019, for Canadian dollar to Euros and American dollar to Euros.

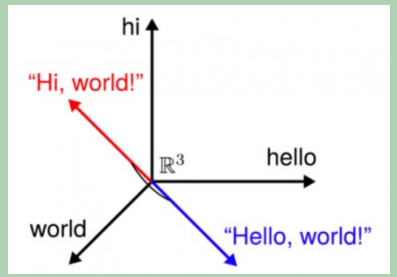
QUERY:

I would like to know the annual average exchange rate for the year 2018, for Canadian $ to Euros and American $ to Euros

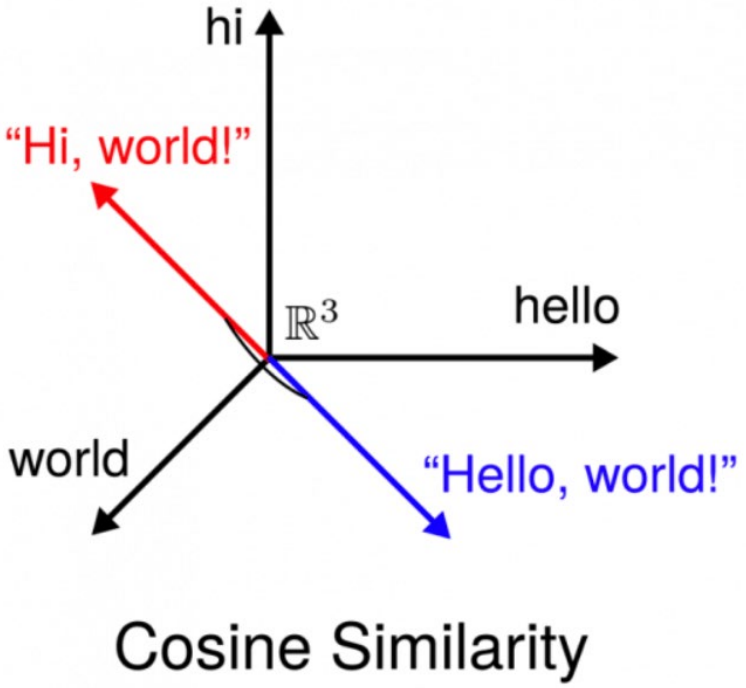
ANSWER:

Please find exchange rates data at the following link.

Cosine similarity



Similarity of queries: base model



$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

Source: machinelearningplus.com, Wikipedia

Application demo