Automated Ticketing System I

Dominik Macko

Situation

- Automate a ticketing system, for new ticket:
  - assign to corresponding team
  - predict time to solve it
  - find similar solved tickets

Situation

- Automate a ticketing system, for new ticket:
  - assign to corresponding team
  - predict time to solve it
  - find similar solved tickets
- Non-public “dirty” anonymized dataset
- ~5000 labeled samples, ~20000 unlabeled samples

ticket	problem_abstract	team	ticket_opened_date	ticket_closed
272431733	The CUC Database is not running on DURPRIPT7B	Prognosis	12/18/2019 14:20	7/27/2022 20:49
271472293	OPEN CR:Alerts not being generated - trap ba...	SMARTS/GFP_CPE	11/20/2019 18:33	12/23/2021 10:44

Tasks

- NLP Tasks – we only have problem abstract text

Tasks

- NLP Tasks – we only have problem abstract text
- Team assignment – multiclass classification

Tasks

- NLP Tasks – we only have problem abstract text
- Team assignment – multiclass classification
- Time to resolve prediction – regression task

Tasks

- NLP Tasks – we only have problem abstract text
- Team assignment – multiclass classification
- Time to resolve prediction – regression task
- Find similar tickets
  - Unsupervised task
  - Can be viewed as a recommender system or information retrieval (new ticket is query)

Proposed Action

- Supervised tasks
  - Maybe design heuristic rule-based model
  - Classical approaches (tfidf followed by ML algorithms)
  - Pretrained word embeddings
  - Pretrained Transformers, maybe few-shot learning for classification

Proposed Action

- Supervised tasks
  - Maybe design heuristic rule-based model
  - Classical approaches (tfidf followed by ML algorithms)
  - Pretrained word embeddings
  - Pretrained Transformers, maybe few-shot learning for classification
- Unsupervised task
  - Classical approaches (tfidf followed by cosine similarity, bm25)
  - Pretrained word embeddings
  - Pretrained Sentence Transformers (Bi-encoder and Cross-encoder)

Existing Solutions

- Many solutions which companies can buy
  - Usually the goal is just to route tickets to correct departments
  - We want solution to our specific data with all our tasks

Existing Solutions

- Many solutions which companies can buy
  - Usually the goal is just to route tickets to correct departments
  - We want solution to our specific data with all our tasks
- In literature
  - Focus on classification
  - Many times metadata used, not just text
  - Often older classical approaches (no neural networks)

Language and Libraries

- Python 3 with Jupyter notebooks
- pandas, numpy, matplotlib, scikit-learn
- gensim, NLTK, rank_bm25
- HuggingFace Transformers, SentenceTransformers, SetFit
- display of results – streamlit or gradio

Evaluation Plan

- Supervised tasks
  - Evaluate on test set using appropriate metrics
  - Imbalanced multiclass classification metrics (F1 micro-averaged)
  - Suitable regression metrics (MSE)

Evaluation Plan

- Supervised tasks
  - Evaluate on test set using appropriate metrics
  - Imbalanced multiclass classification metrics (F1 micro-averaged)
  - Suitable regression metrics (MSE)
- Unsupervised task (ticket similarity)
  - Manually label top N predicted similar tickets for the models
  - Measure MAP@N
  - Manually inspect recall
  - Take into account speed – comparing ticket to all other tickets can’t be too slow!

Thank you for your attention!