A Comparison of Classical and Modern Information Retrieval Approaches on Recipes

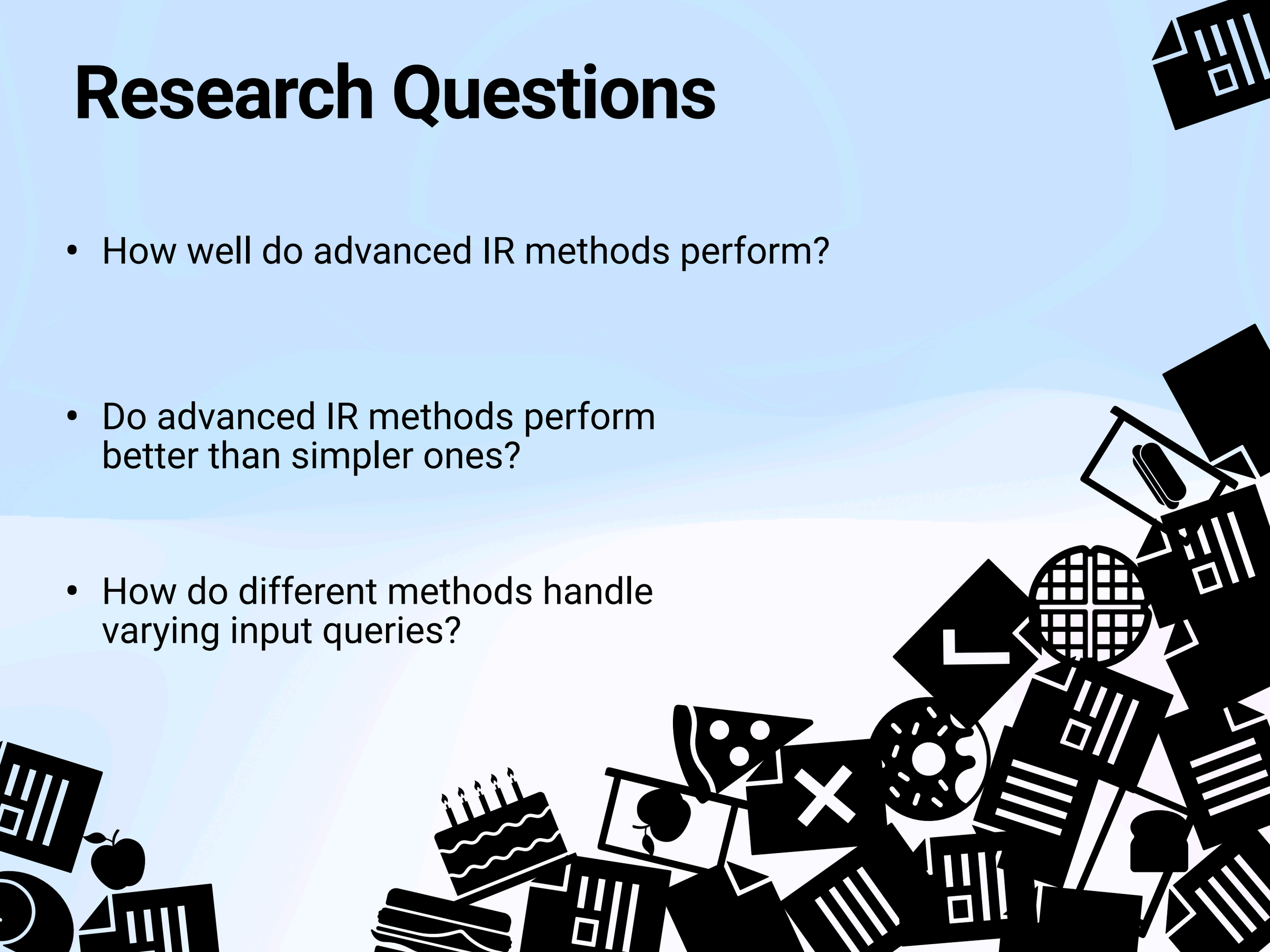
https://github.com/Dowakiin/Advanced-Information-Retrival-WS24-25-Group22

A Project by Group 22:

- Markus Auer-Jammerbund — Team-role: Query pipeline setup, Evaluation, Report
- Thomas Knoll — Team-role: Design Document, *BERT* embeddings, Evaluation, Report
- Jonas Pfisterer — Team-role: *Word2Vec* embeddings, Evaluation, Report
- Thomas Puchleitner — Team-role: Design Document, *TF-IDF* embeddings, Result processing, Evaluation, Report

Research Questions

- How well do advanced IR methods perform?
- Do advanced IR methods perform better than simpler ones?
- How do different methods handle varying input queries?



Data

Where did we get it from and what did we do with it?

Data:

- The recipes dataset and the pre-trained BERT model are from Hugging Face

Pre-processing:

- Merged the ingredient columns
- Merged the instruction columns
- Reduced the size to 100,000 (from 2 mil.)

Models

In increasing complexity

TF-IDF

- The old reliable
- Lightweight
- Easy to understand
- Simple to compute
- Lacks context awareness

Word2Vec

- Newer approach
- Simple in theory
- Not so simple to compute
- Has limited context awareness

BERT

- State-of-the-art
- Transformer based
- Heavyweight
- Has context awareness

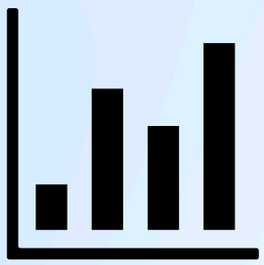
Experiments and Evaluation

Experiment set-up:

- 3 difficulties: easy, medium, and hard
- 3 queries per difficulty
- 7 recommendations per query

Evaluation set-up:

- 4 evaluators
- 63 recommendations per method
- Evaluate for relevance (Yes / No)

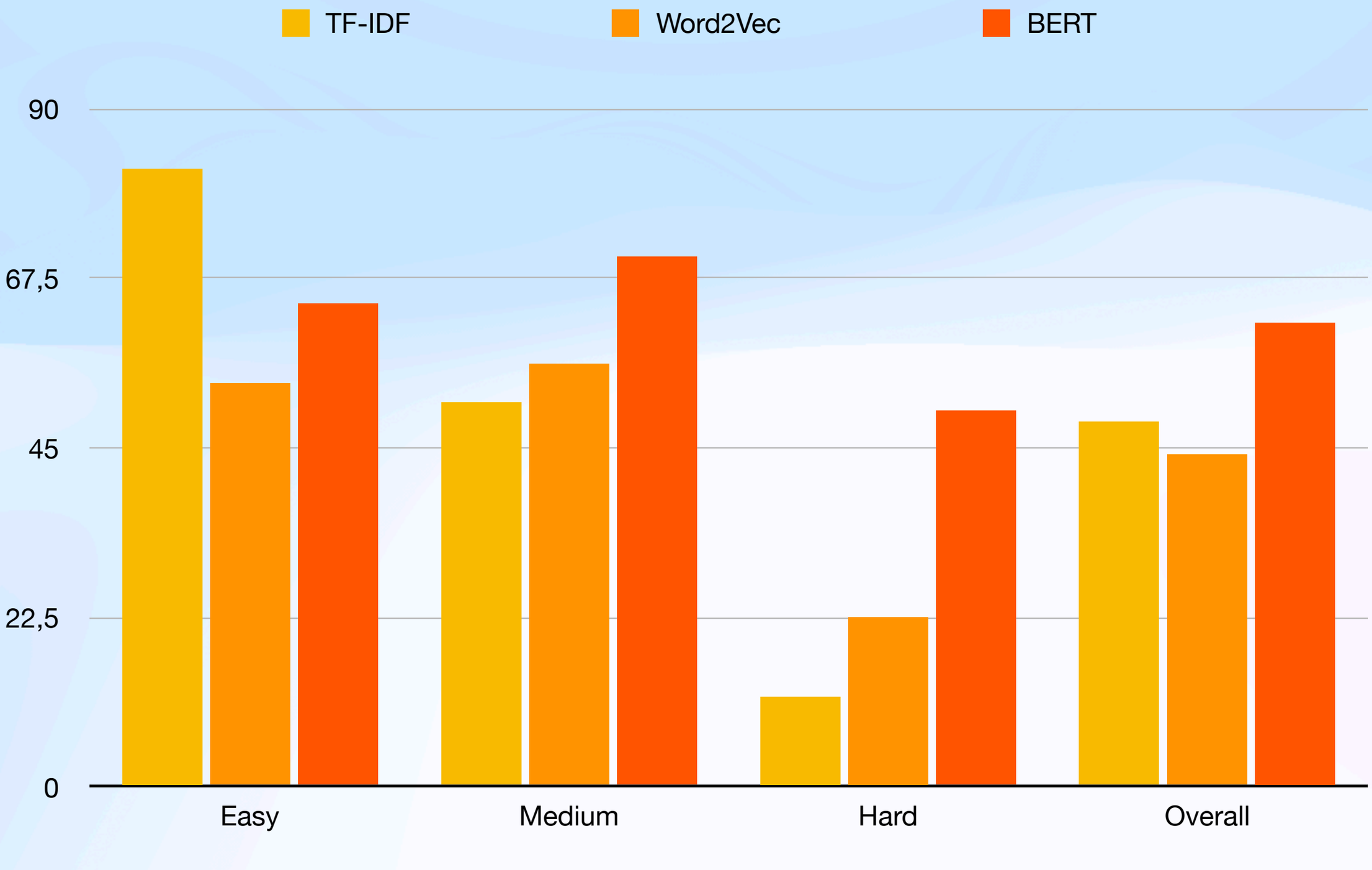


Evaluation:

- Compute the percentage of relevances across evaluators
- Average results to get a score

Results

Recipe relevance per IR method for different query types



Conclusion / Possible future work

- Test with the whole dataset

- Increase robustness by: doing a blind evaluation, asking more queries, utilizing additional evaluators, and expanding on the returned recipes

- Deeper analysis of the weaker performance of *Word2Vec* (focusing on its parameter settings, training conditions, etc.)

- Additional training or fine-tuning of the BERT model