BERT-FlowDelta Improvement and NL2SQL Integration



Project Value: Our project improves a Q&A system model with access to Ford vehicle manuals that can answer customers questions about vehicle features and operation while capturing elements of human conversation

Mario Flores | Brendan Kerr Pum Khai | Brandon Wu

BERT-FlowDelta Model and Improvement

What is BERT-FlowDelta?

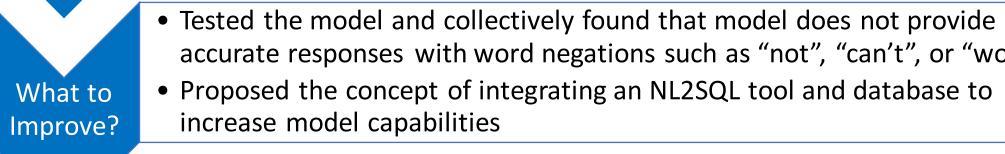
- A multiturn question and answer model that captures human conversation more accurately than single turn models
 - The flow of human conversation (i.e. how topics change) is captured by passing the difference of two previous hidden states
- The Modular Annotation team, built a tool to produce data to further fine-tune the model in the future

Model Improvements

- Originally, when provided with user-generated questions containing word negations, the model provided nonsensical and inaccurate responses
- To improve this, we gave the model user-generated questions that contained negations, validated model answer output, and formatted the answer output to be used as training data

Improvement Process

- Code was readily accessible online with multiple published papers discussing the effectiveness of the model
- Relatively high rank on both CoQa and QuAC Datasets compared to other NLP models



Flow Delta?

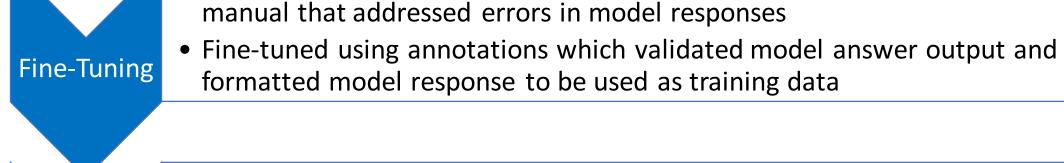
increase model capabilities

accurate responses with word negations such as "not", "can't", or "won't"

• Created question and answer pairs (annotations) from Lincoln Aviator Ford

manual that addressed errors in model responses

formatted model response to be used as training data



- Fine-tuning increased the model response accuracy (F1 score) of 66.4 using
- Expanded model capabilities by integrated NL2SQL to provide model responses from vehicle manual tables

Ford Data to an overall score of 70.7 in the fine-tuned model

Business Value

Integrating NL2SQL with BERT-FlowDelta

NL2SQL

Motivation:

- Ability to generate answers from the tables inside the Lincoln Aviator Manual
- Requirements:
 - Database: A collection of 176 tables extracted from the Lincoln Aviator Manual
 - Metadata table with entities that reference all other tables by a unique identifier (uuid)
 - Natural Language question
 - The corresponding SQL query (generated by X-SQL model)

Updated Question/Answer System Pipeline



Information

Question

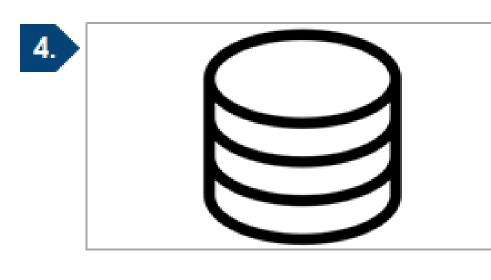
A user generates a question, and that question gets fed into the IR system.

IR System

The IR system uses encoding to narrow down the owner manual by returning only the most relevant paragraphs.

BERT-FlowDelta

Each unique context is concatenated to a larger context.



5. X-SQL



NL Question What safety restraint would you recommend for a small child?



Database

Using the chapter number, section number, and start and end token indices, the system searches for a table in the database, returning the table identification attribute. If no table gets returned, the system defaults to the answer generated by BFD.

X-SQL

If a table identification is returned, the X-SQL model generates an SQL statement based on the question input which then will query the information from that table.

Answer

Information that is returned by the SQL query becomes the newly updated answer outputted to the user. If X-SQL was not used, the IR system will highlight the part of the Ford Owner Manual text that was used as the answer from BFD.

Model Improvement Results and Conclusions

Fine-tuning Results

- Assigned a score to the performance of the model questionresponse accuracy (F1 score).
- Base model (3 epochs)
 - o Ford F1: 36.7
 - Overall F1: 56.4
- Fine-tuned model (3 epochs):
 - o Ford F1: 66.4
 - Overall F1: 70.7
- Ford 20 -10 Base Model Fine-Tuned Model Overall

Fine-Tuned Model

Conclusion

• The F1 score of the base model (pre-training data), fine-tuned model, and overall scores shows that the model was considerably improved due to the data produced by our team. In addition, the NL2SQL integration will help the model answer customers' questions better by gathering information from both the main text and the tables inside the Lincoln Aviator Manual.

Future Goals

- Merge the Modular **Annotation Tool with** BERT-FlowDelta
- SQL Query rejection model integration

Acknowledgements

Thank you, Sarah Rodenbeck, Justin Gould, the Ford Machine Learning group, Dr. Mark Daniel Ward, Ellen Gundlach, and Maggie Betz for providing the resources and support for the successful completion of our project.