UBER

Customer Obsession Ticket Assistant

Improving Uber Customer Support with Natural Language Processing and Deep Learning

Piero Molino | AI Labs
Huaixiu Zheng | Applied Machine Learning
Yi-Chia Wang | Applied Machine Learning

Main Takeaways

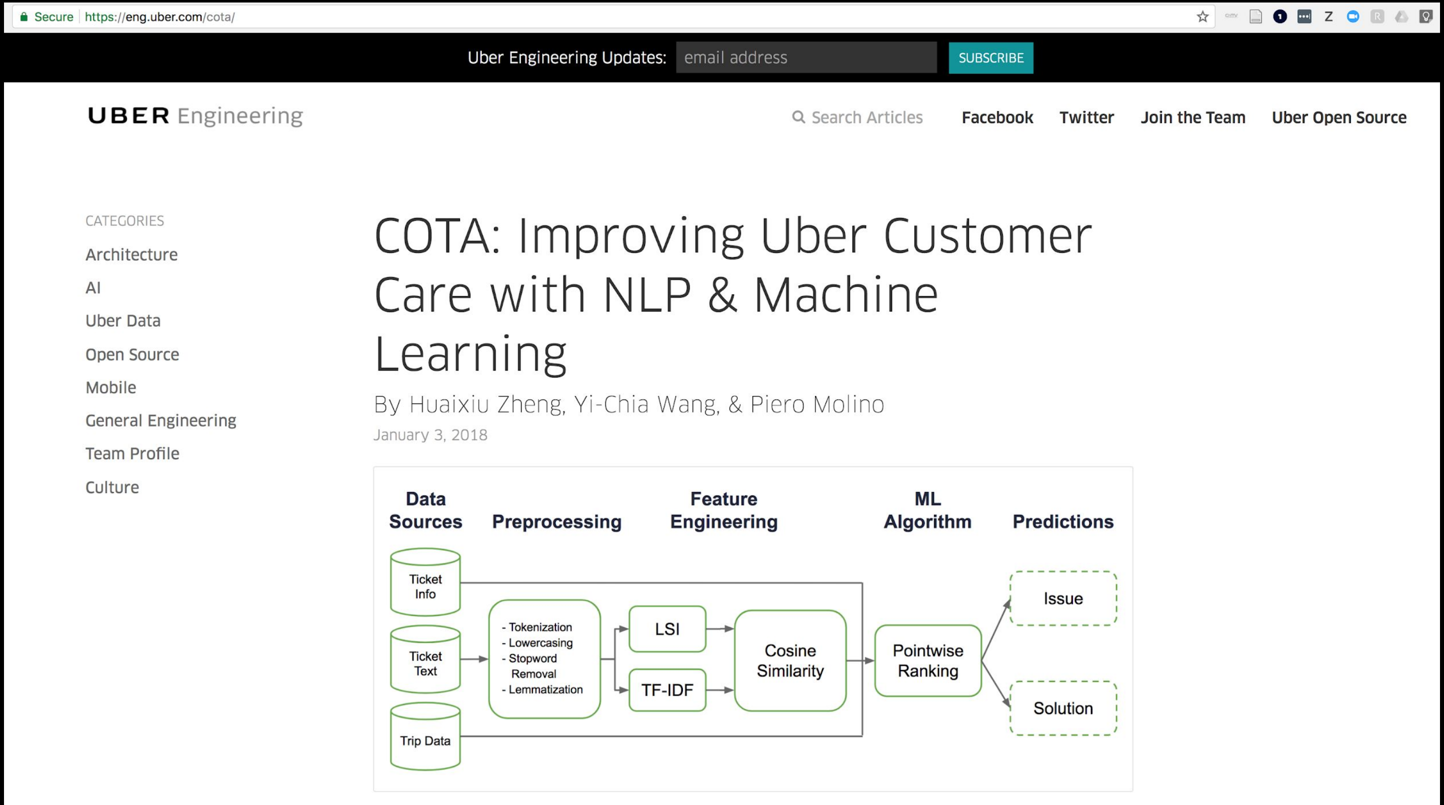
COTA v1: classical NLP + ML models

- Faster and more accurate customer care experience
- Million $ of saving while retaining customer satisfaction

COTA v2: deep learning models

- Experiments with various deep learning architectures
- 20-30% performance boost compared to classical models

COTA Blog Post and followup, KDD paper



Agenda

Motivation and Solution

- Complexity of Customer support @Uber

COTA v1: Traditional ML / NLP Models

- Multi-class Classification vs Ranking

COTA v2: Deep Learning Models

- Deep learning architectures

COTA v1 vs COTA v2

Agenda

Motivation and Solution

- **Complexity of Customer support @Uber**

COTA v1: Traditional ML / NLP Models

- Multi-class Classification vs Ranking

COTA v2: Deep Learning Models

- Deep learning architectures

COTA v1 vs COTA v2

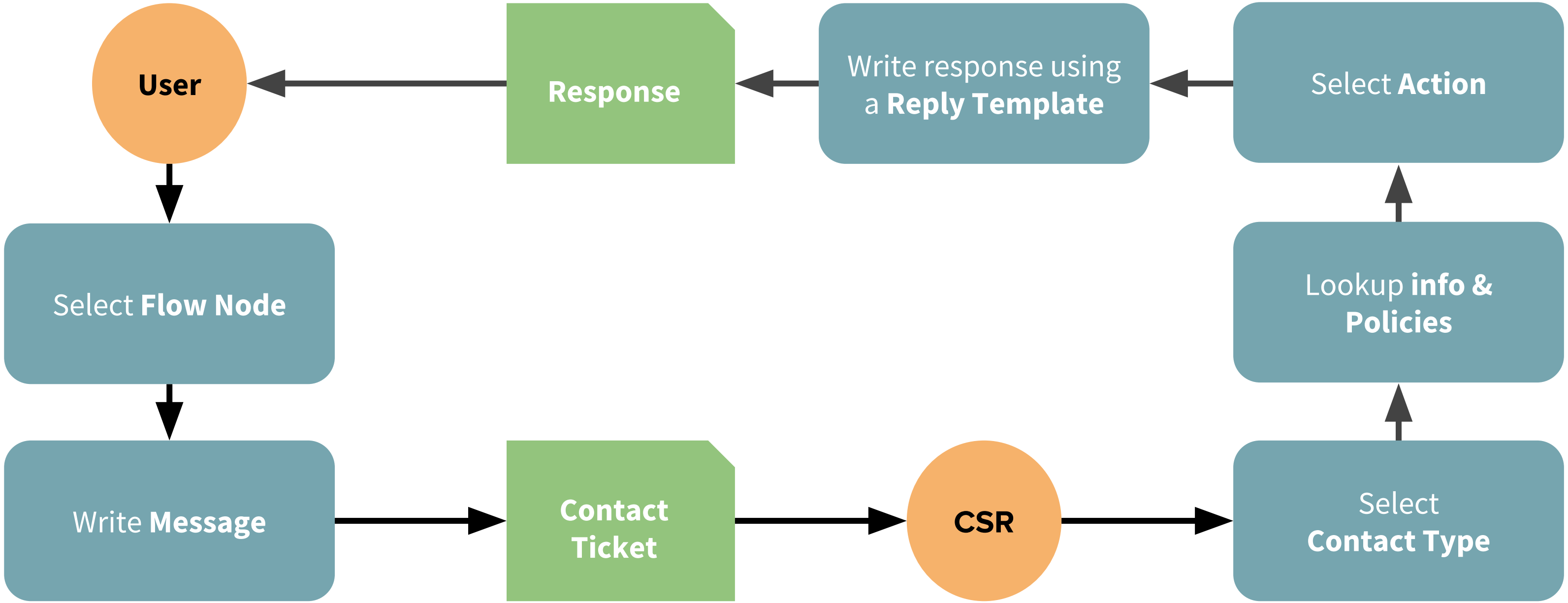
What is the challenge?

As Uber grows, so does our volume of support tickets

Millions of tickets from riders / drivers / eaters **per week**

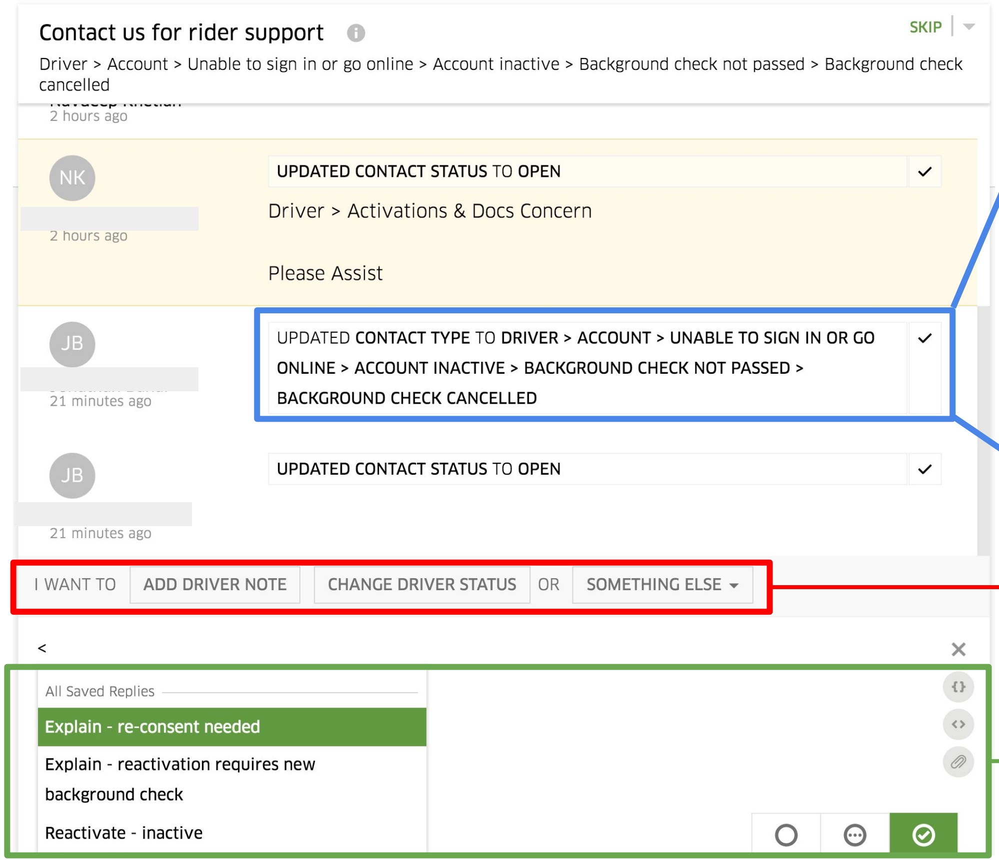
Thousands of different types of issues users may encounter

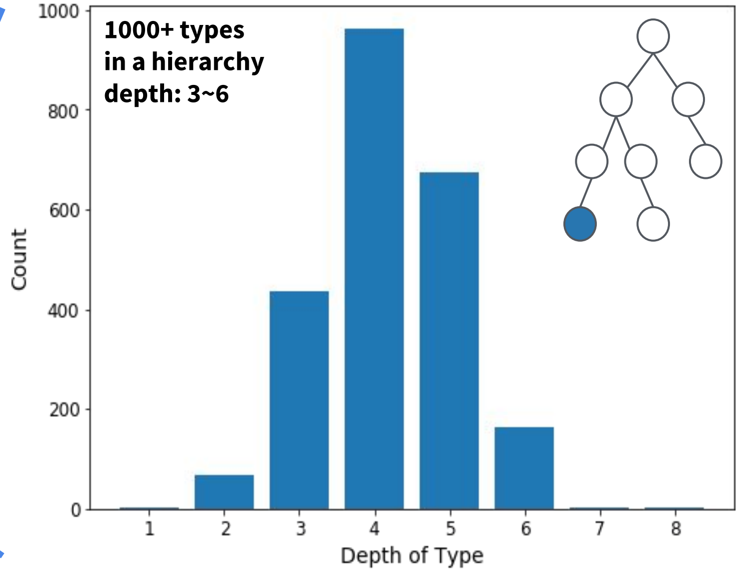
Uber Support Platform



What is the challenge?

And it is not easy to solve a ticket





10+ actions (adjust fare, add appeasement, …)

1000+ reply templates

Agenda

Motivation and Solution

- Complexity of Customer support @Uber

COTA v1: Traditional ML / NLP Models

- **Multi-class Classification vs Ranking**

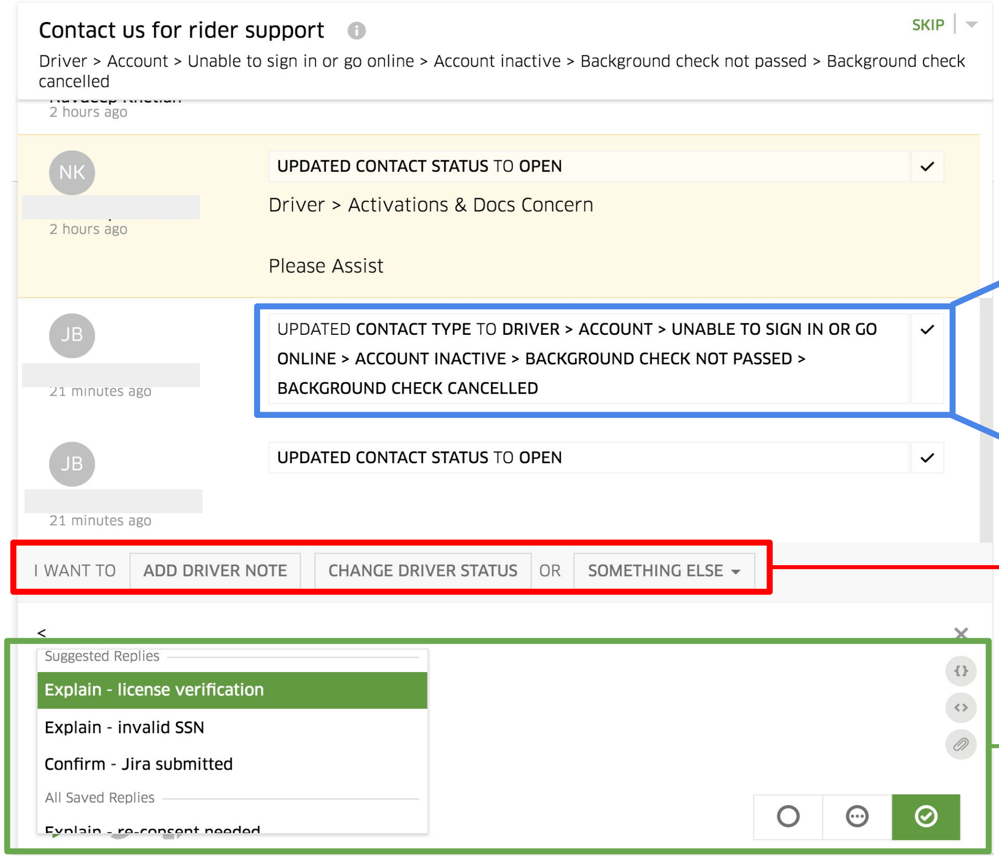
COTA v2: Deep Learning Models

- Deep learning architectures

COTA v1 vs COTA v2

COTA v1: Suggested Resolution

Machine learning models recommending the 3 most relevant solutions



SUGGESTED CONTACT TYPES

Driver > Account > Unable to sign in or go online > Account inactive

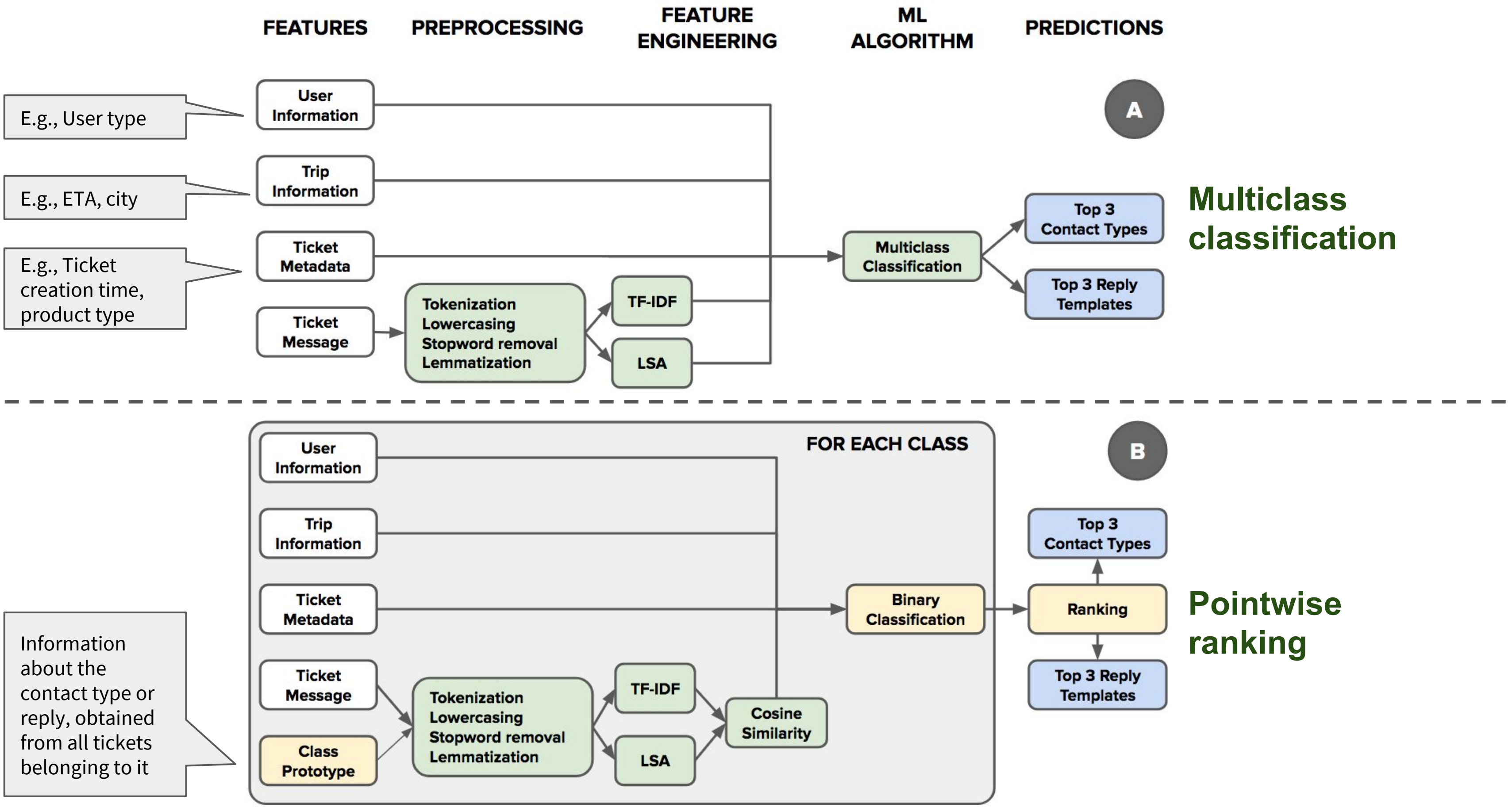
Driver > Account > Profile > Unsubscribe > SMS or Text

Driver > Account > Vehicles > Edit vehicle class

Reorder actions in relevance

Surface top-3 most-relevant reply templates

COTA v1 Model Pipeline



From Classification to Ranking

Multi-class Classification

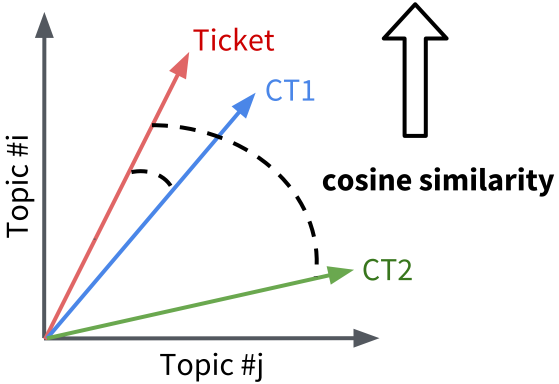
Tickets Features	Label (CT1, CT2)
t1 features	CT1
t2 features	CT2

Pointwise Ranking

Tickets Features	Type Features	Sim (t, CT)	Label (0, 1)
t1 features	CT1 features	0.8	1
t1 features	CT2 features	0.1	0
t2 features	CT1 features	0.2	0
t2 features	CT2 features	0.7	1

Ranking allows us to include **features of candidate types** and **similarity features** between a ticket and a candidate type

Model: **Random Forest** with hyperparameters optimized through **grid search**



Performance Comparison

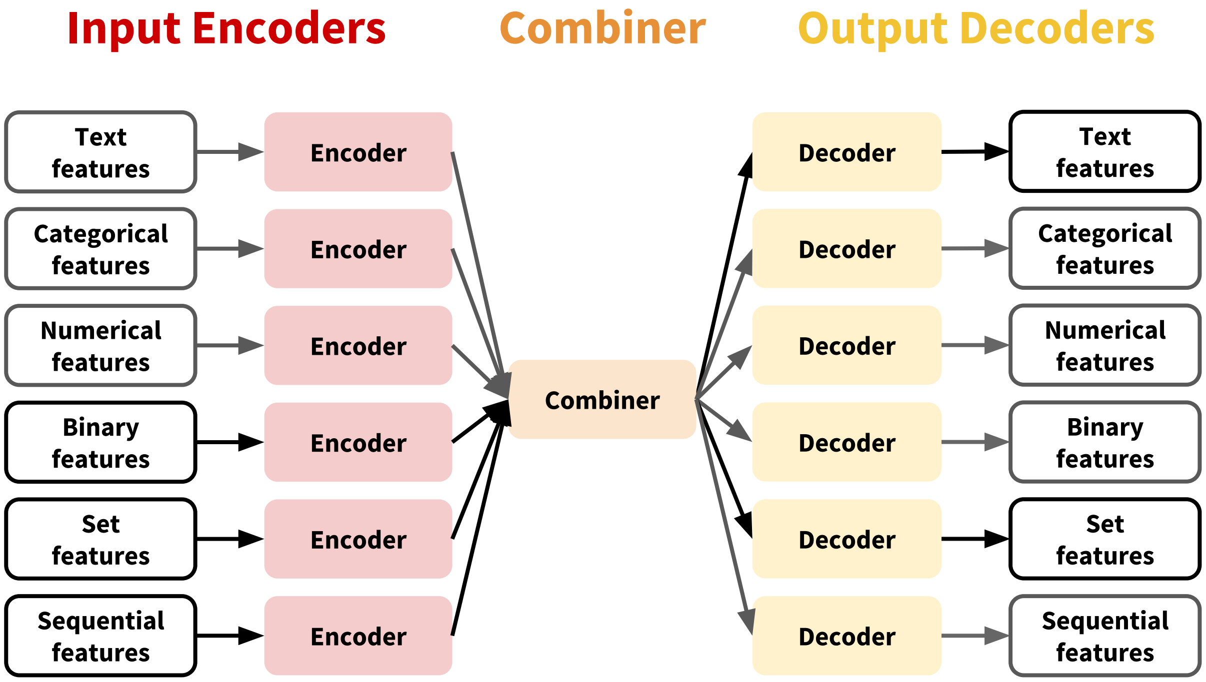
6% absolute (10% relative) improvement

Hits@3: any of the top 3 suggestions is selected by CSRs

Agenda

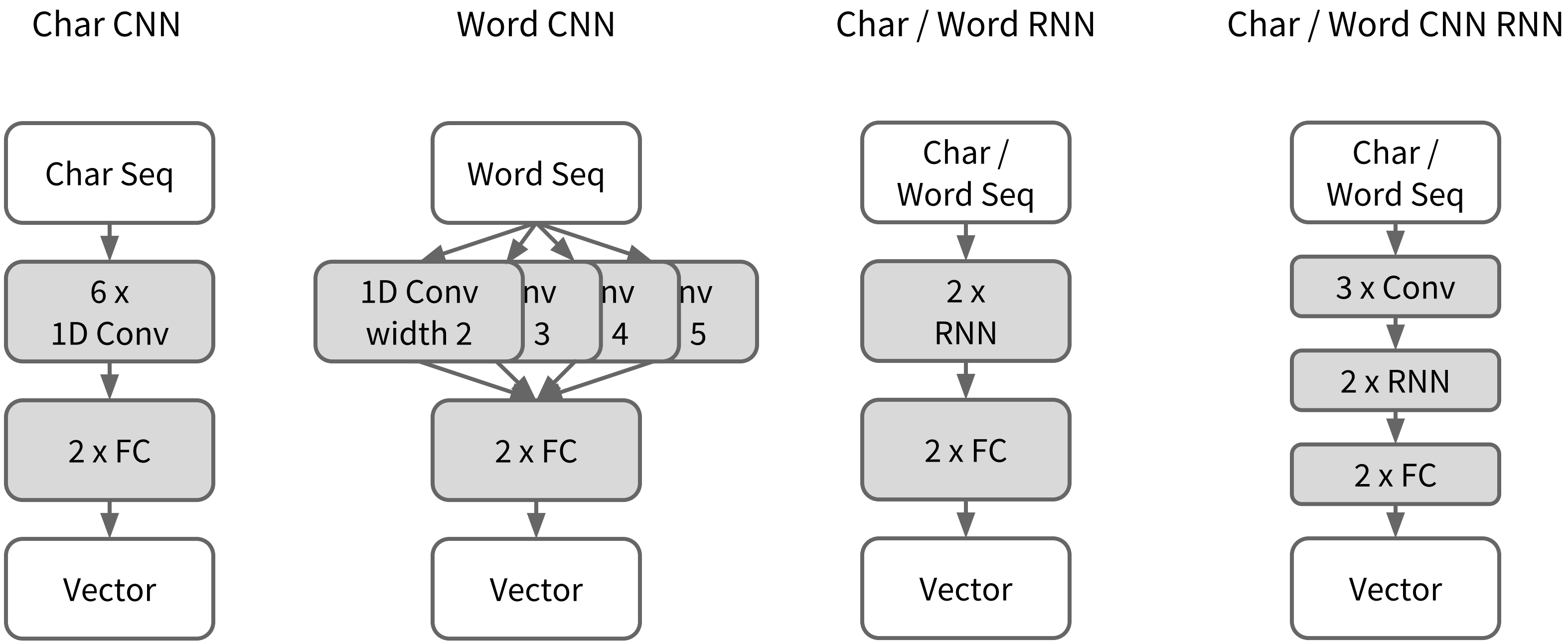
- Motivation and Solution
  - Complexity of Customer support @Uber
- COTA v1: Traditional ML / NLP Models
  - Multi-class Classification vs Ranking
- **COTA v2: Deep Learning Models**
  - **Deep learning architectures**
- COTA v1 vs COTA v2

COTA v2: Deep Learning Architecture



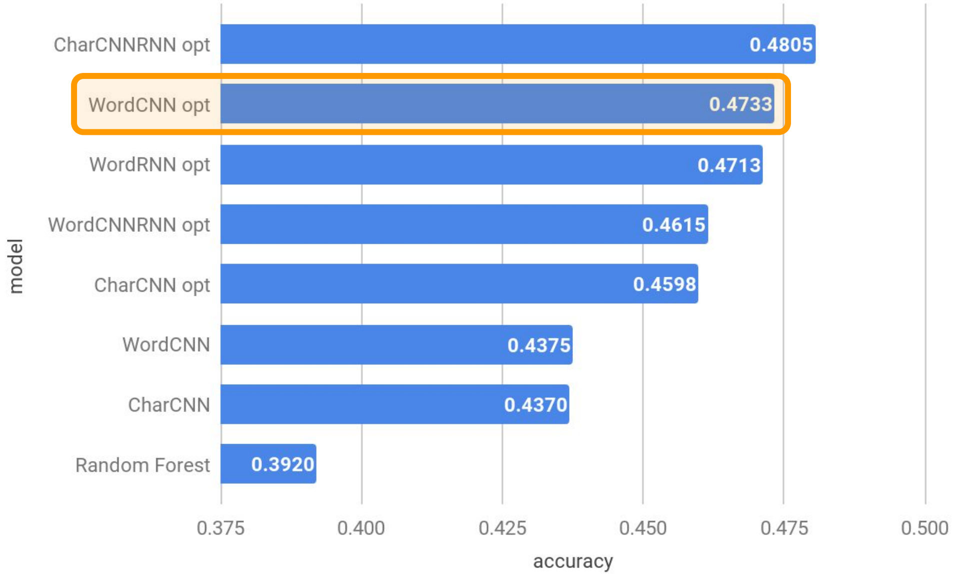
Generic architecture, **reusable** in many different applications.
We are considering open-sourcing it!

COTA v2: Text Encoding Models



Which text encoder?

Hyperparameter search for contact type classification



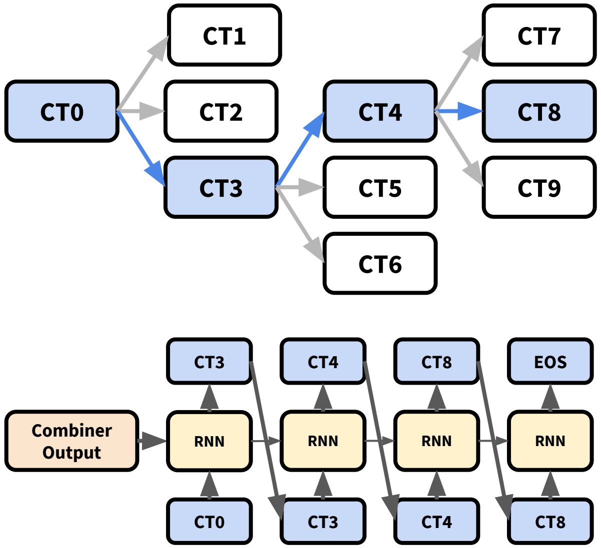
Model	Validation accuracy	Training time per epoch in minutes
CharCNNRNN opt	0.4805	35
WordCNN opt	**0.4733**	**4**
WordRNN opt	0.4713	17
WordCNNRNN opt	0.4615	12
CharCNN opt	0.4598	5

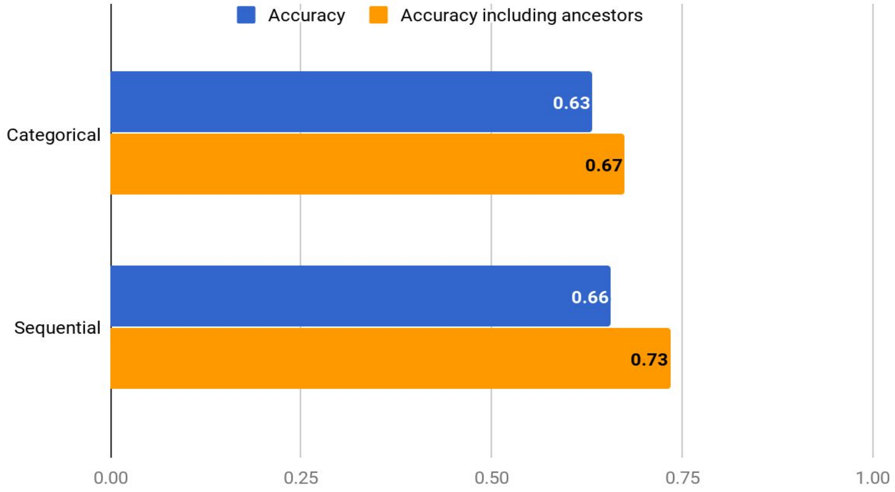
WordCNN is the **best compromise** between **performance** and **speed**

20%+ over Random Forest used in COTA v1 and **~10x** faster than CharCNNRNN

Sequence Model for Type Selection

Predict the sequence of nodes instead of leaf node





Example: **Driver** > **Trips** > **Pickup and drop-off issues** > **Cancellation Fee** > **Driver Cancelled**

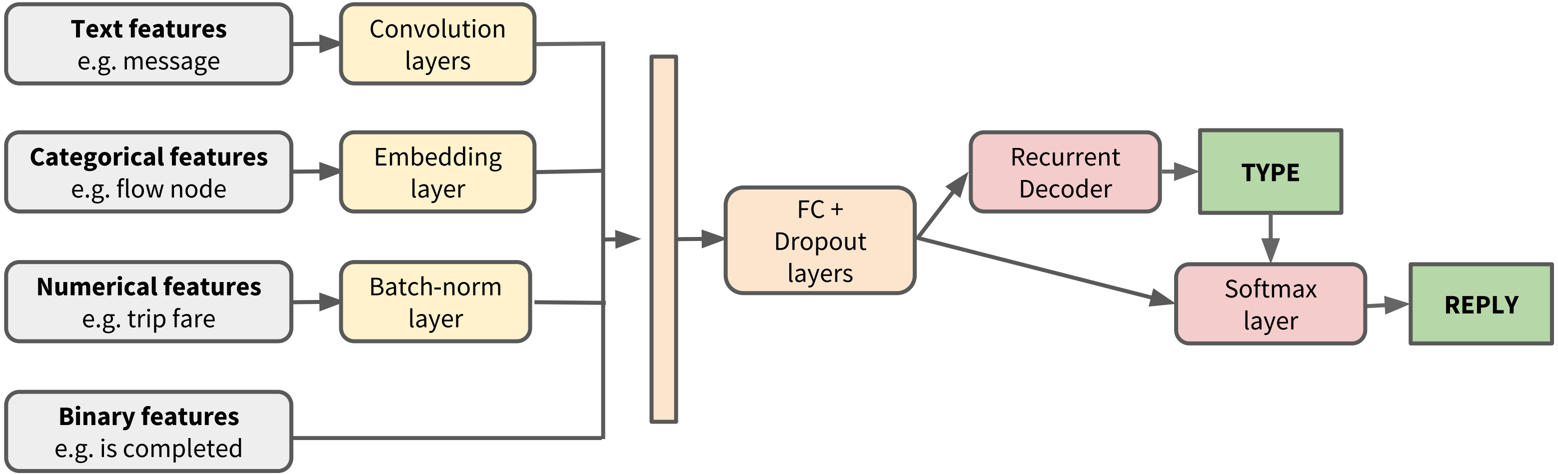
Use a Recurrent Decoder to predict **sequences of nodes** in the contact type tree

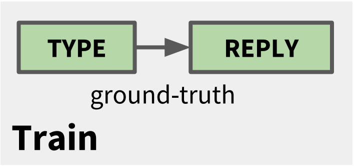
Pick the last class before <eos> as prediction

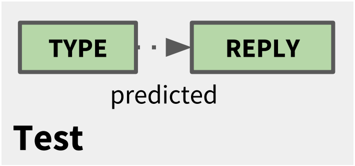
Model makes **more reasonable mistakes**

Final Architecture

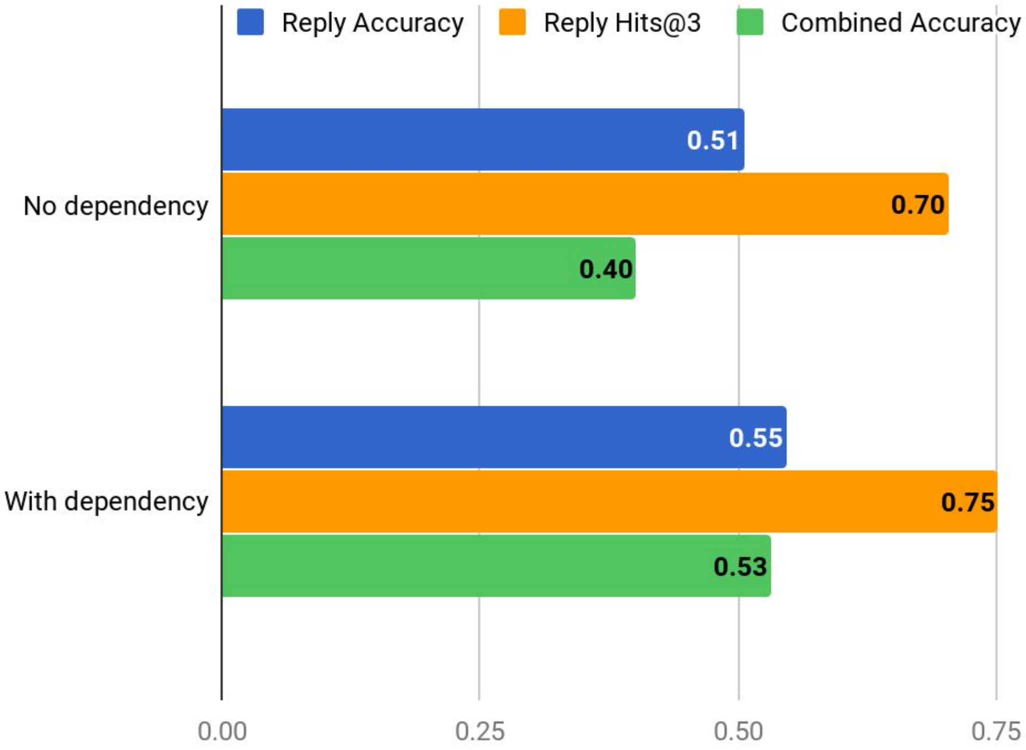
Multi-task sequential learning







Effect of Adding Dependencies Between Tasks



Adding the dependency from Type to Reply **improves accuracy**

It also improves a lot the **coherence** between the two models, **increasing combined accuracy** consistently

Combined accuracy computed requiring both Type and Reply model to be **correct at the same time**

Outline

Motivation and Solution

- Complexity of Customer support @Uber

COTA v1: Traditional ML / NLP Models

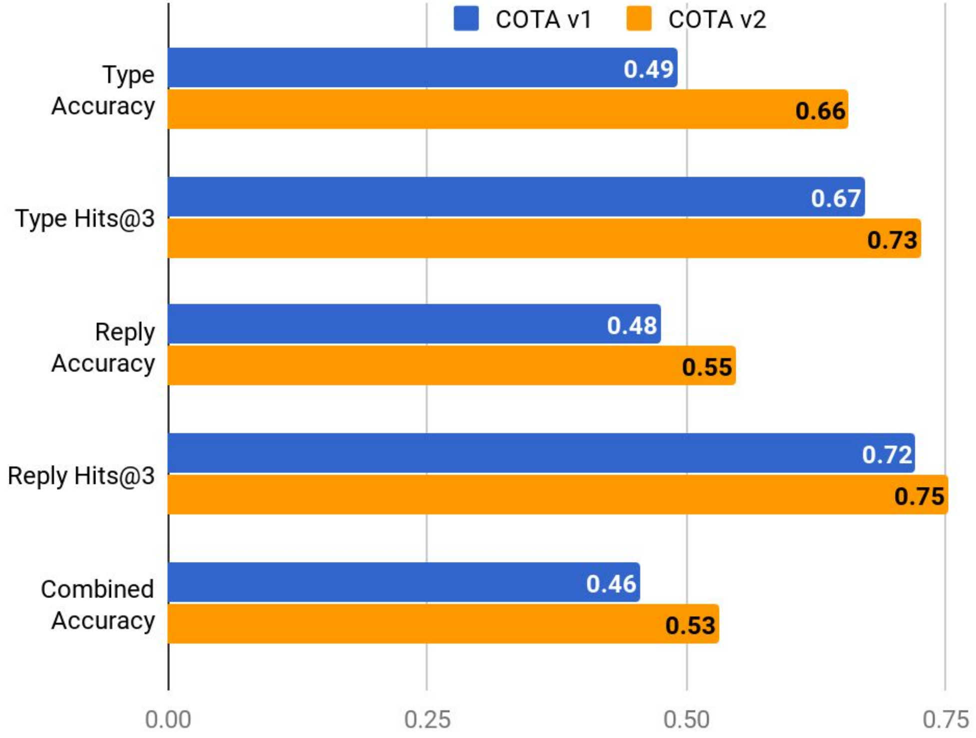
- Multi-class Classification vs Ranking

COTA v2: Deep Learning Models

- Deep learning architectures

COTA v1 vs COTA v2

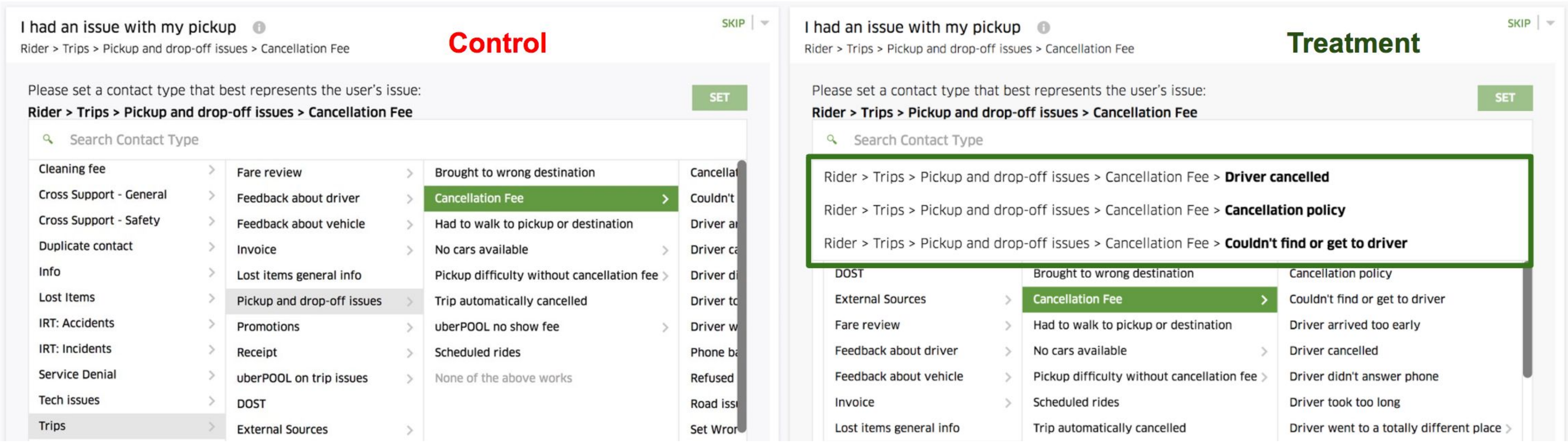
COTA v1 vs. COTA v2 offline comparison



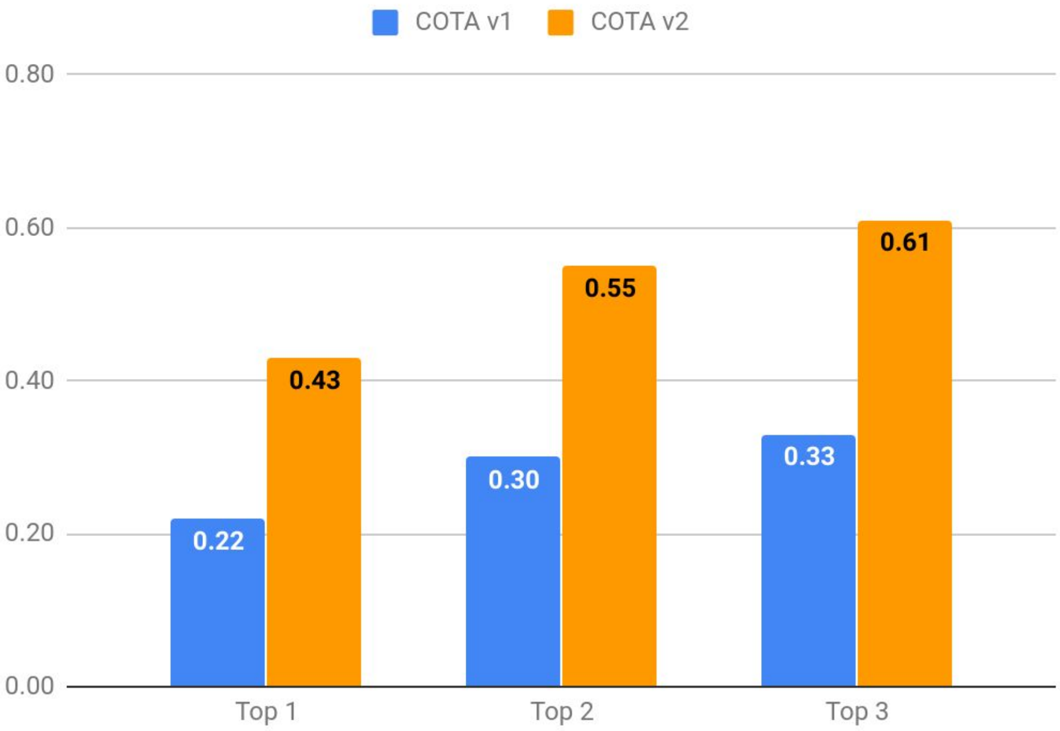
COTA v2 is **consistently more effective** than COTA v1 on **all metrics** for **both models**

The combined accuracy in particular shows an absolute **~+9%** (relative **+~20%**)

COTA v1 vs. COTA v2 A/B Test



COTA v1 vs. COTA v2 A/B Test

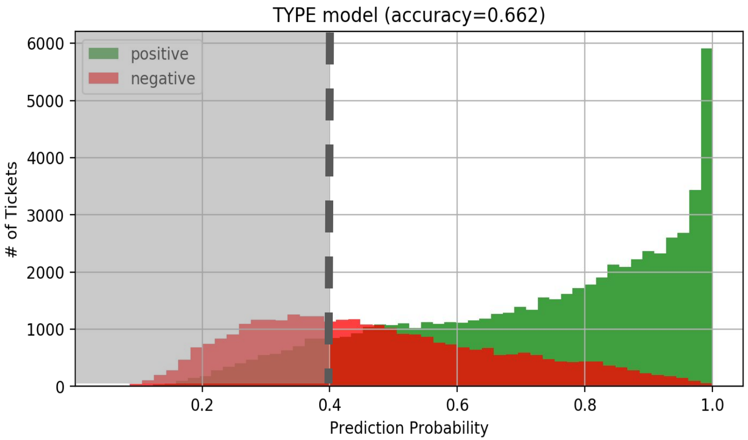


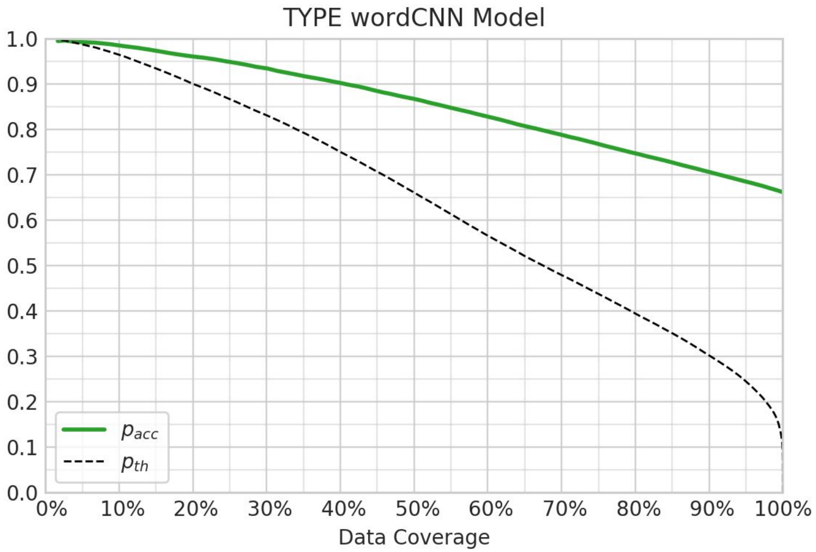
COTA v2 is **20-30% more accurate** than COTA v1 in online A/B tests

COTA v1 **reduces handling time** of **~8%,** while COTA v2 provides an additional **~7% reduction**, more than ~**15% overall reduction**

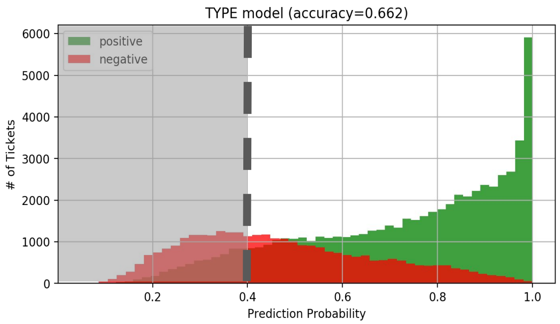
Statistically significant **customer satisfaction improvement**

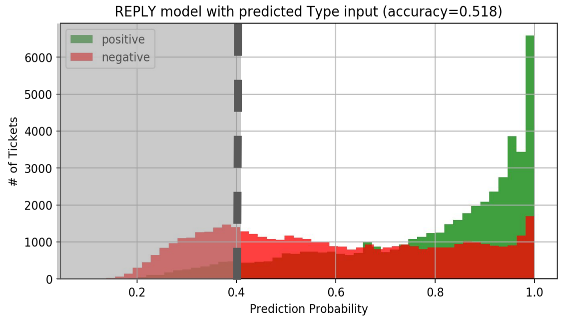
Threshold on Type Model Confidence

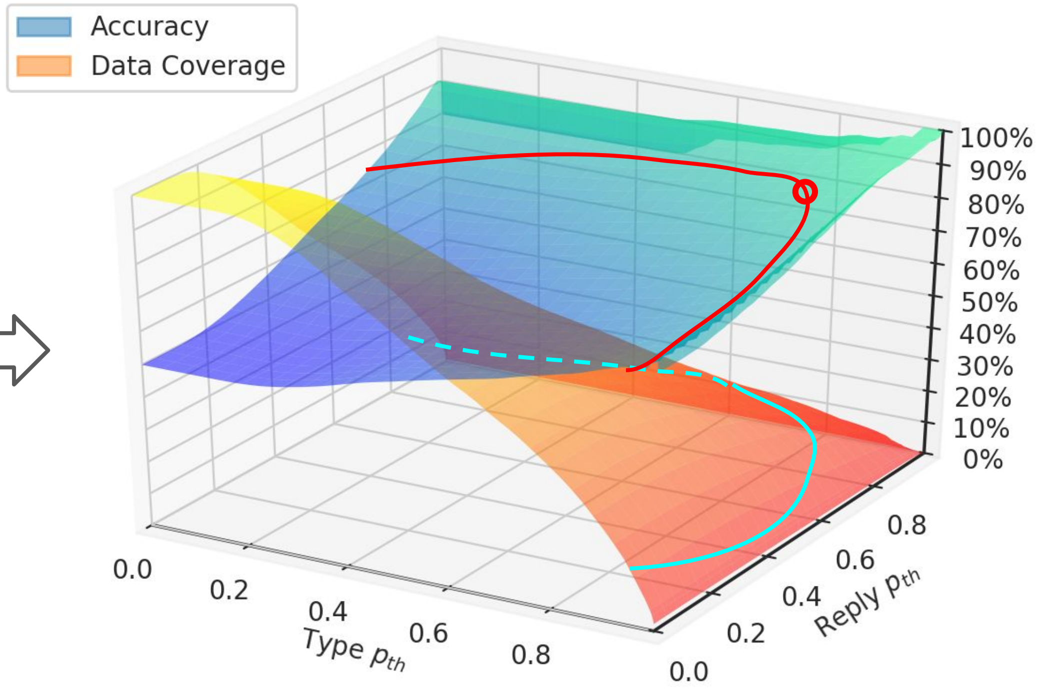




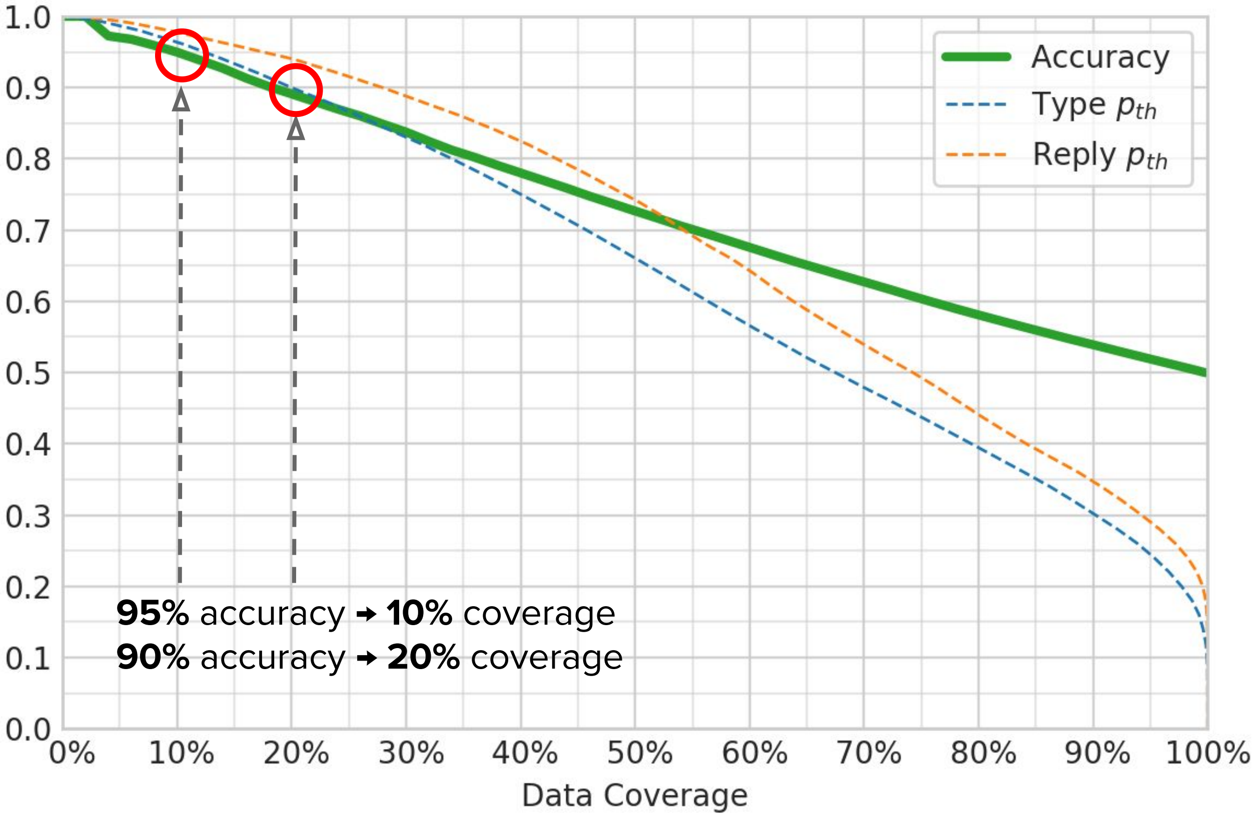
Threshold on Both Models' Confidence







Coverage vs. Maximum Accuracy



Conclusions

Using NLP & ML COTA makes customer care experience **faster** and **more accurate** while **saving Uber millions of $**

Moving from traditional to deep learning models, we observe a substantial **performance boost** (up to **30%**)

Using intelligent suggestions we were able to **reduce ticket handling time without impacting customer satisfaction**

COTA Team

Cross-functional collaboration

AI Labs
Applied Machine Learning
Customer Obsession
Michelangelo
Sensing and Perception

Proprietary and confidential © 2018 Uber Technologies, Inc. All rights reserved. No part of this document may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage or retrieval systems, without permission in writing from Uber. This document is intended only for the use of the individual or entity to whom it is addressed and contains information that is privileged, confidential or otherwise exempt from disclosure under applicable law. All recipients of this document are notified that the information contained herein includes proprietary and confidential information of Uber, and recipient may not make use of, disseminate, or in any way disclose this document or any of the enclosed information to any person other than employees of addressee to the extent necessary for consultations with authorized personnel of Uber.