P11: Kaggle: Riiid! Answer Correctness Prediction

Robert Pikmets, Anna-Helena Salurand, Taavi Kivisik
Problem: Student Answer Correctness Prediction

- Time series of interactions
- Lectures & questions
- Predict per user, probability of the next question is answered correctly
- Code challenge!
Method/approach

- Add models
- Add features
- Tune hyperparameters
Planned and Did

Planned

- **Data exploration**
- Data cleaning
- Add baseline models
- Feature engineering
- Model upgrading and tuning
- Final presentation prep

Did

- 101M user interactions (answers, lectures)
- 393K unique users
- 66% of all answers correct
Planned and Did

**Planned**
- Data exploration
- **Data cleaning**
- Add baseline models
- Feature engineering
- Model upgrading and tuning
- Final presentation prep

**Did**
- Handled (few) missing data
- Handled types (for memory!)
  - e.g. float64 to float32
  - object to bool
# Planned and Did

**Planned**
- Data exploration
- Data cleaning
- **Add baseline models**
- Feature engineering
- Model upgrading and tuning
- Final presentation prep

**Did**
- XGBoost with sample 1, AUC 0.5
- + failed attempts due to memory usage
Planned and Did

Planned

- Data exploration
- Data cleaning
- Add baseline models
- **Feature engineering**
- Model upgrading and tuning
- Final presentation prep

Did

- Added question correctness std: WORSE!
- Failed attempts due to memory limits:
  - → per-question average correctness
  - → per part average correctness
<table>
<thead>
<tr>
<th>Planned</th>
<th>Did</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data exploration</td>
<td></td>
</tr>
<tr>
<td>Data cleaning</td>
<td></td>
</tr>
<tr>
<td>Add baseline models</td>
<td></td>
</tr>
<tr>
<td>Feature engineering</td>
<td></td>
</tr>
<tr>
<td><strong>Model upgrading and tuning</strong></td>
<td>Hyperparameter tuning:</td>
</tr>
<tr>
<td>Final presentation prep</td>
<td>→ e.g. decreasing learning rate from 0.1</td>
</tr>
<tr>
<td></td>
<td>to 0.05, higher AUC</td>
</tr>
</tbody>
</table>
Planned and Did

Planned

● Data exploration
● Data cleaning
● Add baseline models
● Feature engineering
● Model upgrading and tuning
● ++
● Final presentation prep

Did

● Reduced training size for memory
## Planned and Did

<table>
<thead>
<tr>
<th>Planned</th>
<th>Did</th>
</tr>
</thead>
<tbody>
<tr>
<td>● Data exploration</td>
<td></td>
</tr>
<tr>
<td>● Data cleaning</td>
<td></td>
</tr>
<tr>
<td>● Add baseline models</td>
<td></td>
</tr>
<tr>
<td>● Feature engineering</td>
<td></td>
</tr>
<tr>
<td>● Model upgrading and tuning</td>
<td></td>
</tr>
<tr>
<td>● ++</td>
<td>✓</td>
</tr>
<tr>
<td>● Final presentation prep</td>
<td></td>
</tr>
</tbody>
</table>
Results

- **AUC score**
  - Intermediate presentation 0.745 (10M training data)
  - → now worse 0.712 (1M training data)
Main lessons

- Competition restrictions matter - memory limits
- Don’t read other notebooks too often
Main lessons

- Competition restrictions matter - memory limits
- Don’t read other notebooks too often
- Don’t leave it too late
- Model training - time-consuming
Main lessons

- Competition restrictions matter - memory limits
- Don’t read other notebooks too often
- Don’t leave it too late
- Model training - time-consuming

“Due to the inference API and the time/memory constraints, this competition is really as much about software engineering as it is about ML.” – Kaggle grandmaster (Silogram) on the competition
Code sharing

- Due to Kaggle competition rules we are allowed to only share code outside the team by making our notebook public. Link:
The End

Thank you