**RESULTS – Model Development**

<table>
<thead>
<tr>
<th>Feature Window</th>
<th>ROC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 session</td>
<td>0.6814</td>
<td>0.8855</td>
<td>0.1447</td>
<td>73.45%</td>
</tr>
<tr>
<td>2 days</td>
<td>0.6814</td>
<td>0.8855</td>
<td>0.1447</td>
<td>73.45%</td>
</tr>
<tr>
<td>3 days</td>
<td>0.7056</td>
<td>0.9014</td>
<td>0.1957</td>
<td>74.78%</td>
</tr>
<tr>
<td>7 days</td>
<td>0.7174</td>
<td>0.9142</td>
<td>0.2334</td>
<td>75.43%</td>
</tr>
</tbody>
</table>

Above models are built using Logistic Regression.

• A zero rule classifier (always chooses majority class) has 70.86% accuracy (ROC = 0.5).

• For one week feature window, considering only activity count as predictor improves the accuracy by 5.43% and ROC by 0.21.

• ASCORE and DSCORE when considered together, can improve the model accuracy by 6.06% and ROC goes up by 0.22.

• Adding session count & time between sessions to the model with ASCORE & DSCORE improves the accuracy by ~10% and ROC by 0.28.

**RESULTS – Model Comparison**

<table>
<thead>
<tr>
<th>Feature Window</th>
<th>ROC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 session</td>
<td>0.6822</td>
<td>0.8952</td>
<td>0.1908</td>
<td>74.11%</td>
</tr>
<tr>
<td>2 days</td>
<td>0.6822</td>
<td>0.8952</td>
<td>0.1908</td>
<td>74.11%</td>
</tr>
<tr>
<td>3 days</td>
<td>0.7098</td>
<td>0.9142</td>
<td>0.2020</td>
<td>75.57%</td>
</tr>
<tr>
<td>7 days</td>
<td>0.7274</td>
<td>0.9226</td>
<td>0.2527</td>
<td>76.04%</td>
</tr>
</tbody>
</table>

• Gradient Boosting models consistently have high accuracy and ROC across all feature windows.

• Logistic Regression models have high specificity whereas Naive Bayes models have high sensitivity across most of the feature windows.

• Random forest models have good balance of sensitivity and specificity but perform marginally poor compared to GB.

**CONCLUSION**

• Lower the number of diverse activities in the first week, higher is the probability to churn.

• More widespread the sessions are in the first week, lower is the probability of churn; more the number of sessions in the first week, lower is the probability of churn.

• ASCORE and DSCORE remain significant predictors in explaining churn across all feature windows.

• The overall accuracies and ROC’s of Gradient Boosting models exceed those of other 3 machine learning models. However Logistic Regression and Naive Bayes models are individually better in terms of specificity and sensitivity respectively.

**ACKNOWLEDGEMENT**

I would specially like to thank Prof. Joseph Konstan and Raghav Karunmur for their constant guidance and support throughout the project. I would also like to thank Prof. De Liu and Prof. Singhandhu Chatterjee for their valuable feedback on the project.

**REFERENCES**


---

**METHODOLOGY**

Model Features:
- Session count
- Time between sessions
- ASCORE
- DSCORE

DSCORE for an user is the degree of dissimilarity of distinct activity types performed by the user in certain time interval based on the hierarchical ontological relatedness of these activity types.

Tree diagram on the left depicts the relationship between different activity types on Movielens.

For instance,
- 3 users having activity pattern as \((w, w, e, y, z, l, w, y, v, z, x, z, w, x, y, z)\) will have same set of activities \((w, x, y, z, l)\).
- If these activities are represented by below tree diagram & distance matrix \(d\) then, \(DSCORE = \sum_{i=1}^{n} w_{ij}\)