ppAURORA: Privacy Preserving Area Under Receiver Operating Characteristic and Precision-Recall Curves

Ali Burak Ünal, Nico Pfeifer, Mete Akgün
Motivation

● Data is everywhere!
Motivation

- Data is everywhere!
- Machine learning algorithms demand data.
Motivation

- Data is everywhere!
- Machine learning algorithms demand data.
- Privacy of the sensitive information!
Motivation

- Data is everywhere!
- Machine learning algorithms demand data.
- Privacy of the sensitive information!
  - Privacy preserving model training and testing

![Diagram showing model training and testing process with data sets and model variables](attachment:image.png)
Motivation

- Data is everywhere!
- Machine learning algorithms demand data.
- Privacy of the sensitive information!
  - Privacy preserving model training and testing
  - How about the privacy preserving model evaluation such as the area under curve?

\[
\begin{align*}
M_{\text{initial}} & \rightarrow M_{\text{trained}} & \hat{y} &= \text{AUC}(y, \hat{y}) \\
D_{\text{train}} & \rightarrow D_{\text{test}} & 0.91 & 0.23 & 0.89 & 0.76 & 0.11 & \cdot & \cdot & 0.87
\end{align*}
\]
ppAURORA

- Privacy preserving model evaluation based on 3-party computation (MPC) framework\textsuperscript{[1]}

ppAURORA

- Privacy preserving model evaluation based on 3-party computation (MPC) framework[1]

ppAURORA

- Privacy preserving model evaluation based on 3-party computation (MPC) framework\textsuperscript{[1]}

ppAURORA

- Privacy preserving model evaluation based on 3-party computation (MPC) framework\[^1\]

ppAURORA

- Privacy preserving model evaluation based on 3-party computation (MPC) framework\cite{unal2022ceclia}

ppAURORA

- Privacy preserving model evaluation based on 3-party computation (MPC) framework

- Area under the curve (AUC) as the model evaluation metric
  - Summarizes the plot-based model evaluation metrics by measuring the area between the curve and the x-axis
  - Receiver operating characteristic (ROC) curve
  - Precision-Recall (PR) Curve
ppAURORA

- Privacy preserving model evaluation based on 3-party computation (MPC) framework

- Area under the curve (AUC) as the model evaluation metric
  - Summarizes the plot-based model evaluation metrics by measuring the area between the curve and the x-axis
  - Receiver operating characteristic (ROC) curve
  - Precision-Recall (PR) Curve

- Exact AUC computation via the MPC building blocks
  - Especially for the small size test set
Area Under the ROC Curve (AUROC)

- ppAURORA for the area under the ROC curve (AUROC)
Area Under the ROC Curve (AUROC)

- ppAURORA for the area under the ROC curve (AUROC)

- Two versions
  - No tie condition in the prediction scores (AUROC no-tie)
  - With tie condition in the prediction scores (AUROC with-tie)
Area Under the ROC Curve (AUROC)

- For AUROC no-tie

\[ AUROC = \frac{\sum_{i=1}^{M} \left( TP[i] \cdot (FP[i] - FP[i-1]) \right)}{T \cdot F} \]
Area Under the ROC Curve (AUROC)

- For AUROC no-tie

\[
AUROC = \frac{\sum_{i=1}^{M} (TP[i] \cdot (FP[i] - FP[i-1]))}{T \cdot F}
\]

- # true positives
- # false positives
- # true samples
- # false samples
Area Under the ROC Curve (AUROC)

* For AUROC no-tie

\[
AUROC = \frac{\sum_{i=1}^{M} (TP[i] \cdot (FP[i] - FP[i-1]))}{T \cdot F}
\]

- \(TP[i]\): the number of true positives
- \(FP[i] - FP[i-1]\): the number of false positives
- \(T \cdot F\): the total number of samples
### Why AUROC *with-tie*?

<table>
<thead>
<tr>
<th>Prediction Score</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
</tr>
</tbody>
</table>
Why AUROC with-tie?

<table>
<thead>
<tr>
<th>TPR</th>
<th>FPR</th>
<th>Prediction Score</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.4</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.6</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.8</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.6</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.8</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
</tr>
</tbody>
</table>

Graph showing True Positive Rate vs. False Positive Rate.
Why AUROC with-tie?

<table>
<thead>
<tr>
<th>TPR</th>
<th>FPR</th>
<th>Prediction Score</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.4</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.6</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>0.8</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.6</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.8</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
</tr>
</tbody>
</table>

![ROC Curve](image)
Area Under the ROC Curve (AUROC)

- For AUROC no-tie

\[
AUROC = \frac{\sum_{i=1}^{M} \left( TP[i] \cdot (FP[i] - FP[i-1]) \right)}{T \cdot F}
\]

- For AUROC with-tie

\[
AUROC = \sum_{i=1}^{\Theta} \left( \frac{(TP[i] + TP[i-1]) \cdot (FP[i] - FP[i-1])}{2 \cdot T \cdot F} \right)
\]
Area Under the ROC Curve (AUROC)

- For AUROC no-tie

\[
AUROC = \frac{\sum_{i=1}^{M} (TP[i] \cdot (FP[i] - FP[i-1]))}{T \cdot F}
\]

- For AUROC with-tie

\[
AUROC = \sum_{i=1}^{\Theta} \left( \frac{(TP[i] + TP[i-1]) \cdot (FP[i] - FP[i-1])}{2 \cdot T \cdot F} \right)
\]

where: all samples, # true positives, # false positives, # true samples, # false samples, threshold samples determined via secure tie detection.
Area Under the ROC Curve (AUROC)

- For AUROC no-tie

\[
AUROC = \frac{\sum_{i=1}^{M} (TP[i] \cdot (FP[i] - FP[i-1]))}{T \cdot F}
\]

- For AUROC with-tie

threshold samples determined via secure tie detection

\[
AUROC = \sum_{i=1}^{\Theta} \left( \frac{(TP[i] + TP[i-1]) \cdot (FP[i] - FP[i-1])}{2 \cdot T \cdot F} \right)
\]
Area Under the Precision-Recall Curve (AUPR)

● ppAURORA for the area under the PR curve (AUPR)
Area Under the Precision-Recall Curve (AUPR)

- ppAURORA for the area under the PR curve (AUPR)

- Similar to AUROC with-tie
  - Precision and recall can change at the same time.
  - No common denominator though
Area Under the Precision-Recall Curve (AUPR)

- ppAURORA for the area under the PR curve (AUPR)

- Similar to AUROC with-tie
  - Precision and recall can change at the same time.
  - No common denominator though

\[
AUROC = \sum_{i=1}^{n} \left( PRE[i-1] \cdot (REC[i] - REC[i-1]) + \frac{(PRE[i] - PRE[i-1]) \cdot (REC[i] - REC[i-1])}{2} \right)
\]
**Area Under the Precision-Recall Curve (AUPR)**

- **ppAURORA** for the area under the PR curve (AUPR)

- **Similar to AUROC with-tie**
  - Precision and recall can change at the same time.
  - No common denominator though

\[
AUROC = \sum_{i=1}^{\Theta} \left( PRE[i-1] \cdot (REC[i] - REC[i-1]) + \frac{(PRE[i] - PRE[i-1]) \cdot (REC[i] - REC[i-1])}{2} \right)
\]
Sorting

- The first task to perform before both AUROC and AUPR
  - Individually sorted lists from multiple data sources

- Merging individually sorted lists using the MPC building blocks
  - Parametric sorting algorithm adjusting the privacy-performance trade-off

- Skipping due to the time limitation
Results

- **Correctness analysis on**
  - Acute Myeloid Leukemia dataset
  - UCI Heart Disease dataset
  - Same as the result of the plaintext analysis

- **Scalability analysis on**
  - Synthetic dataset
  - Various scenarios
Results: Scalability to the Number of Samples
Results: Scalability to the Number of Parties

![Graph showing scalability to the number of parties with different metrics: AUPR, AUROC no-tie, and AUROC with-tie. The x-axis represents the number of parties, and the y-axis represents time in seconds. As the number of parties increases, the time required for the tasks also increases.]
Results: Scalability to the Delta
Summary

- Not only the training and testing privately but also evaluation privately
- ppAURORA based on 3-party computation for AUC of ROC and PR curves
- Exact AUC result
- Linearly scalable to the number of samples and the parties
- Logarithmic decrease in the execution time parallel to the increase in delta
Thanks for listening!

Any Questions?

The icons in this presentation are from https://www.flaticon.com/