Trustworthiness Re-use of Pre-trained Neural Networks (ICSE’23)

Wenxin Jiang, Taylor R. Schorlemmer, James C. Davis

Motivation

The software engineers we interviewed broadly followed four steps.
1. Identify a task and determine if reuse is appropriate.
2. Select an architecture and a particular PTM.
3. Apply reuse techniques and evaluate model performance — a potentially repetitive stage.
4. Deploy the models.

Methodology

Qualitative Study

Research Questions

1. Interviews (N=12)
2. Helpful Attributes (RQ2)
3. Reuse Challenges (RQ3)
4. Risk Mitigation (RQ4)

Quantitative Study

H²Torrent

Risk Mitigation Measurement

Results

Figure 3. Diagram of the four-stage decision making model for PTM reuse.

The software engineers we interviewed broadly followed four steps.
1. Identify a task and determine if reuse is appropriate.
2. Select an architecture and a particular PTM.
3. Apply reuse techniques and evaluate model performance — a potentially repetitive stage.
4. Deploy the models.

Figure 4. Proportion of models with standardized machine-readable performance claims. Organized by PTM task type.

Implications

- Model audit: large-scale measurements
- PTM standardization: Information extraction, ONNX
- Adversarial attack detection: detection tools

PTMTorrent Dataset for Mining Open-source Pre-trained Model Packages (MSR’23)

Data Collection

PTM Registries

GitHub repo

Model Zoo

PyTorch Hub

Model Hub

Data Schema

Table 1. Number of models and storage size for each collected model registry.

<table>
<thead>
<tr>
<th>Name</th>
<th># Models</th>
<th>Data Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hugging Face</td>
<td>12,401</td>
<td>61TB</td>
</tr>
<tr>
<td>Model Zoo</td>
<td>3,245</td>
<td>115GB</td>
</tr>
<tr>
<td>PyTorch Hub</td>
<td>49</td>
<td>1.5GB</td>
</tr>
<tr>
<td>ONNX Model Zoo</td>
<td>185</td>
<td>441MB</td>
</tr>
<tr>
<td>Modelhub</td>
<td>33</td>
<td>721MB</td>
</tr>
<tr>
<td>PTMTorrent</td>
<td>15,913</td>
<td>--</td>
</tr>
</tbody>
</table>

Figure 6. Data collection and processing workflow for PTMTorrent.

We standardize data and metadata based on a common PTMTorrent data schema.


Future Work

- Supporting future PTM supply chain analysis.
- Expanding PTM model registry analysis.
- Furthering the state of mining tool development.

More Information