A User-Interactive Machine Learning Model for Identifying Structural Relationships of Code Features

By: Andrea Solis, Kartik Chugh, Ankit Gupta
Introduction

Problems with tools using Machine Learning:
- User doesn’t know what is happening
- Get’s not say on how the system works

Applied in a tool that combines both Machine Learning and User Interaction:
- Applied to an HTML editor and Auto-Complete
- Includes User Interface
Introduction

HTML Structural patterns:
- Elements in same section
- Type of element
- Size
- Position
- Color
- Etc.
Code Completion

Types of code completion:
- Tag
- Attribute
- Value

.parent's tag
.parent's attribute-value pair

Tag
Attribute
Value

<body bgcolor="#E6E6FA">
<a href="gmu">George Mason</a></body>
Auto Complete Process

Every time user types something:

- Determines the type of prediction
- Get Abstract Syntax Tree (AST) from source code
Auto Complete Process

- Obtain training data, and sample features from AST.

<table>
<thead>
<tr>
<th>Parent’s Tag</th>
<th>Parent’s Attr-Val Pair</th>
<th>Tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>body</td>
<td>Id=&quot;foo&quot;</td>
<td>(prediction)</td>
</tr>
</tbody>
</table>
Auto Complete Process

- Build Decision Tree.
- Use sample features and get a prediction from Decision Tree.
- Show the predictions as options in Auto Complete.
Results for Auto Complete

- Without any user feedback
- Leave-one-out Validation
- Using 13 websites
User Interface

Information given to the user:
- Pattern rule for current top prediction
- Option to see other pattern rules
User Interface

Changes made by the user:
- Prioritize or blacklist current prediction
- Change current prediction
- Add a new pattern rule
Pilot Study

2 tasks for users:
- Start a document from scratch
- Complete an existing document

Features added to User Interface
- Highlight examples
- Option to unprioritized pattern rules
Future Work

Auto Complete:
- Improve value prediction

User Interface:
- Change the way the rules are shown to the user
- Based on future pilot studies

Testing:
- 2nd Pilot study, formal user study
- More validation tests
Thank you!