Detecting Programming Errors Using Machine Learning

Eisha Ahmed, Joshua Campbell, Andrea McIntosh, Dr. Abram Hindle
Department of Computer Science, University of Alberta

What is Python?

- Python is a widely used programming language used to do tedious and repetitve tasks.
- One small error, such as an extra bracket or missed indentation, can cause the program to break

Terminal Window displaying syntax error

```
  File "cbpy2", line 51
    if code == 1
                ^
SyntaxError: invalid syntax
```

Research Purpose

- Find breaks and errors in other programmers' codes.
- The collected data will be used to find and locate any errors that future programmers might make in their code.

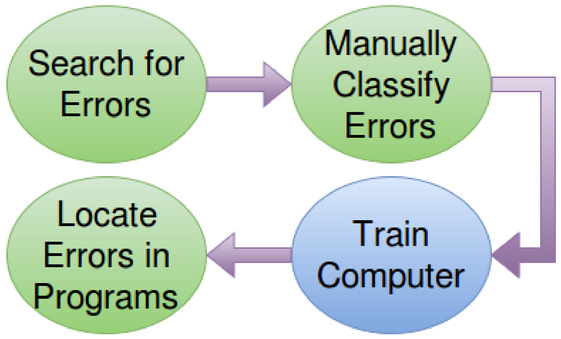
```
    temp_fragment = ''
    if file_system == 'hfs+'
    if file_system == 'hfs+':
        normalize = unicodedata.ucd_3_2_0.normalize
    else
    else:
        normalize = unicodedata.normalize
    for character in sanitized_fragment:
```

Original Text with error

Corrected Version

A Sample Mistake. Forgetting the colon after an if or else statement may result in a break in program

Procedure



Methodology used to train the computer in mistake classification

- Used GitHub to collect sample coding data .
- Train the computer to find errors using Weka

Actual value of data

Prediction by computer

```
1:yes   2:no
1:yes   1:yes
1:yes   2:no
1:yes   1:yes
1:yes   1:yes
1:yes   1:yes
1:yes   1:yes
1:yes   2:no
1:yes   1:yes
1:yes   2:no
1:yes   1:yes
2:no    2:no
2:no    2:no
2:no    2:no
2:no    2:no
2:no    2:no
2:no    2:no
2:no    2:no
```

Examples of classification results from Weka. A prediction is correct if the two instances match: yes is classifies as yes and no is classified as no

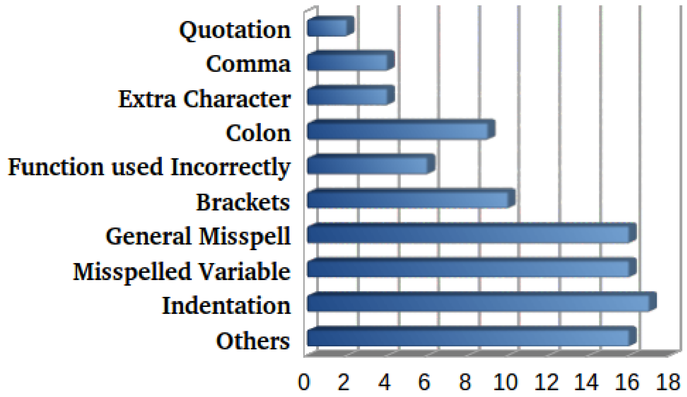
Correct predictions

```
  a   b   <-- classified as
 17  17 |  a = yes
  3  74 |  b = no
```

Incorrect predictions

Results

- Some types of errors occur more frequently than others.



Examples of common types of errors and their frequencies. Generally, Indentation errors happen to occur more frequntly.

- We managed to get our predictions to reach ~80% accuracy. The results can be used to help future programmers find 8 out of 10 errors in their code.

Acknowledgements

- Special thanks to:
  - WISEST and HIP students, Hannah Stormer and Monica Bui.
  - Joshua Campbell for supervising the lab, Andrea McIntosh for her support, and the Principle Investigator Dr. Abram Hindle.
  - Sponsor Canada Summer Jobs
  - WISEST 2016 reasearch program

UNIVERSITY OF ALBERTA WISEST women in scholarship, engineering, science & technology Canada