

1 Introduction: Part 1

1.1 PROBLEM STATEMENT

What problem is your project trying to solve? Use non-technical jargon as much as possible. You may find the Problem Statement Worksheet helpful.

Source code contains many different types of “tags” or words that belong to groups that perform certain functions. With the assistance of machine learning and natural language processing, we hope to create an algorithm that will correctly predict what “tag” any one word in source code belongs to, then, with the help of machine learning, provide a quick comment on what the code is doing. There is a small suite of tools available already, but our client (and PhD advisor) wants to expand these tools. These tools are trained by neural networks that are essentially a “black box” and don’t reveal the connections they make to the engineer. The point of the PhD’s research is to uncover these relationships and see if they are predictable.

1.2 INTENDED USERS AND USES

Who will use the product you create? Who benefits from or will be affected by the results of your project? Who cares that it exists? List as many users or user groups as are relevant to your project. For each user or user group, describe (1) key characteristics (e.g., a persona), (2) need(s) related to the project (e.g., a POV/needs statement), and (3) how they might use or benefit from the product you create. Please include any user research documentation, empathy maps, or other artifacts as appendices.

Users:

Arushi Sharma, PhD student

User details:

Getting their PhD and doing research delving into the depths of machine learning. Dealing with the bleeding edge.

Needs from the project:

Need to be able to see the inside workings of the machine learning to better operate and manipulate the algorithm.

Benefits from the project:

Due to their nature as a PhD student this will help them push machine learning to depths that it has never been before.

Machine learning engineers

User details:

Employed/ salaried engineers with presumably many years in the industry, working on cutting edge technology.

Needs from the project:

Source code, documentation on how we got our data, training and test data, results.

Benefits from the project:

Implementation into their own projects, inspiration for their projects, contribution to the project and valuable insights.

Machine learning enthusiasts/ hobbyists

User details:

Interested in machine learning, most likely a technical background already.

Needs from the project:

Easy-to-read narrative on the project, illustrations, well-documented source code.

Benefits from the project:

Help further understand machine learning, inspiration to begin working on a project, easy to read and digestible information in natural language processing.