**Shalini Panthangi**

**Power Grid Reliability Proposal and Feasibility Study**

**Objective:**

**Purpose:** The purpose of this project is to use various machine learning algorithms which use past environmental and utility data regarding power outages in order to predict future power outage events. Once this prediction is made, the information can be used to decide which power grid systems need to be given more attention to improve so that power supply efficiency can be maximized.

**Expected Outcomes:** The algorithms are expected to identify patterns in the data that is provided and make accurate predictions which will be tested using testing data. The models used should at least give 80% accuracy when predicting future power outage events.

**Justification:**

This research is especially important to utility companies and governments so that they can better plan which grids to pay more attention to in order to improve the efficiency of the power distribution. These kinds of predictions can help citizens and governments better plan their restoration efforts in order to allow people to receive a proper power supply.

**Description/Limitations:**

**Broad explanation of how work will be accomplished:** Data from previous power outage events will be used to formulate various algorithms, including the Support Vector Machine and the Bayesian Additive Regression Tree. These algorithms will be validated using a cross-validation technique, and the final algorithms will be tested for accuracy using a test data set.

**General techniques that will be used:** General techniques that will be used are python coding, formulating the algorithms, and cross-validating the models.

**Variables:** The independent variables in this study are the conditions surrounding a power grid, including outage cause, location, precipitation, soil type, wind, population, etc. The dependent variables in this study are the duration and frequency of the power outage event as predicted by the machine learning algorithm.

**Major constraints:** One major constraint in this study is that no model is perfect, so the model has to be constantly trained and fine-tuned. In addition, if too much data is used, the model will overfit the data, causing the future predictions to fail because the data so perfectly fits past data. Therefore, it is important to not use too much or too little data to train the model.

**Parts of Feasibility Study:**

**What resources are currently available and from where?**

**Personnel:** Available personnel for this project are Mr. Writer to help with the project idea and help with the mathematics and machine learning part of the project. The teachers from the online courses are also available to teach machine learning concepts.

**Equipment:** Currently available equipment is a computer, which is provided by the Academies of Loudoun. This standard laptop is perfectly fine for this research; a larger laptop is not needed.

**Supplies:** Supplies currently available for this project is the data set from the OE-417 power outage events to train the models.

**Knowledge/Skills:** Knowledge and skills already available for this project are python programming skills, basic machine learning algorithm knowledge such as linear algebra and multivariate calculus, and knowledge of the causes of power outages in order to determine key features for the decision tree algorithms which will be used.

**What additional resources are needed?**

**Personnel:** Some additional personnel needed for the research are past researchers who studied the same topic to advise with how to get data or how to train the models with the specific type of data that is going to be used. In addition, the teachers from future machine learning courses will also be needed to teach more complex machine learning concepts.

**Equipment:** Python IDLE will need to be installed in the provided laptop. This will contain all python libraries to create the algorithms.

**Supplies:** Data regarding geologic features of the particular location of the power outage event will be needed for the model to provide a better prediction of the duration and frequency of the power outage event.

**MSDS for chemicals:** No MSDS for chemicals are needed for this project.

**Knowledge/Skills:** Knowledge of cross-validation models, methods to create the machine learning algorithms, and ways to perform accuracy testing are additional skills that are needed for this research. In addition, data sorting skills will be needed for this project in order to carefully organize the data for easy usage.

**Proposed budget:** The proposed budget for this project is no more than 75 dollars, which mainly consists of the possible fees for the machine learning courses.

**Risk assessment:**

**Safety issues and resolution of each:** There are no safety issues for this research except for that the laptop should be handled with care.

**Alternatives explored to achieve you proposed objective:** An alternative that can be done to achieve the objective is to use traditional math modeling approaches with the past data instead of machine learning algorithms. This includes differential equations to find patterns instead of formulating new algorithms for the data. In addition, if some of the models fail, other models can be adopted such as neural network models, random forest regression, etc.