Members:

|  |  |  |
| --- | --- | --- |
| Name | Email | NetID |
| Tao Mo | taomo2@illinois.edu | taomo2 |
| Russell Michal | rmichal2@illinois.edu | rmichal2 |
| Zhenxi Zhou | zzhou36@illinois.edu | zzhou36 |

Project Idea: Explaining Radiation Anomalies with Machine Learning

Abstract:

In the aftermath of the 2011 Fukushima accident, the ability to detect for anomalies in radiation levels received renewed interest. Background radiation is present everywhere, and is known to fluctuate quite significantly with respect to weather conditions, time of day, seasonal variations, and environmental factors. A spike in radiation levels may not be caused by the accidental release of dangerous radioactive isotopes due to a nuclear power plant lost of containment accident or a vehicle carrying medical radioactive isotopes spilling its contents in an accident. The spike may be caused by precipitation or higher than normal humidity, which has been known to cause radiation levels to increase quite substantially above normal background levels due to the release of radioactive radon gas from the ground; caused by precipitation and high humidity. Another natural cause for increase radiation levels is the concentration of natural uranium and thorium in building materials such as granite, which have been known to increase background radiation levels as significant as precipitation and high humidity levels. These radioactive sources are relatively harmless unless a person is exposed to them for very prolonged periods of time in the timescale of decades. In the case of radon gas, the background radiation levels will go back to normal once the precipitation stops and the humidity level decreases.

At UIUC, the RDII research group in the nuclear engineering department is currently in the process of deploying a mobile sensor network comprised of smart phones paired with very small handheld radiation detectors to collect and stream radiation data to the cloud along location, velocity, and timestamp data. For this project, we will analyze this data with machine learning techniques in an attempt to learn if high radiation levels are caused by an environmental factor or actual radiation leak. This will involve using a library such as LB java to apply a set of machine learning techniques to the data. We also have to create several features related to weather and monument location.

We have made features that are set to 0 if the data reading is not near a monument and 1 if the data reading is near a monument. To do this we had to fist get a list of monuments and there GPS coordinates. Then we compared all the monuments coordinates with the data points coordinates. Our first attempt to get monument locations was to use the factual data base that had all the information in it. This way when we could later scale the project to a large scale without much effort. Unfortunately, it would have cost 5000$ to gain access to the data base. Instead we decided to use a google maps like application that had several waypoints listed, sorted by various labels including, parks, restaurants, monuments, and more.

With the location of the monuments and the data point, we were able to create the feature. At first we made the feature 1 if the data point was within 36ft of a monument. This approach did not work because some monuments emit more radiation then others and so, they need a larger radius. Also, some monuments such as old buildings are very large and a 36ft radius would not even cover the whole building. While other monuments like the Alma Matter are small and a 36ft radius would be too big. To get around this problem we gave each monument its own radius based on how far away the farthest data point that had a high radiation level was. This approach worked well for monuments that emitted radiation but it did not work for monuments that did not emit radiation. For example, a grave yard can be labeled as a monument on our map application but will not emit any radiation. For these monuments, we set the distance to a fixed amount like we had done at first.

After creating several features based on whether we were close to a monument or not we then created one feature based on the weather. To make the weather feature we found an API that would give use the weather at the time the data point was measured. We had some problems finding an API that would give use free historical weather information and when we did, the API limited the number of calls we could make. To get around this we fist looked through all our data and only asked the API for weather data from a day if it was in our data. We also did not ask for the same day twice. Unfortunately, the API would tell us that it was raining when our sensors did not measure any high radiation. This was because people tend to stay inside when it’s raining and our sensors are on people’s phones. This problem made this feature unusable because the feature ended up being unnecessary noise.

We still need to find a way to make a weather feature. Possibly by looking at our data points and if the radiation level is higher for one day compared to the others then we could use that to set the weather feature. We will do this in the future. We also need to run several machine learning algorithms on the data and analysis the results. We plan to run our features through decision trees, SVM, Winnow, and perceptron. Possibly more algorithms will be used time permitting.

Papers:

ClariSense+: An Enhanced Traffic Anomaly Explanation Service Using Social Network Feeds,

[http://www.sciencedirect.com/science/article/pii/S1574119216000444](https://webmail.illinois.edu/owa/redir.aspx?C=jROXlVsn5p2v1HgpUfWh0NIwEXeOSlcPmYzNFhc-4VEXEPgyFGjUCA..&URL=http%3a%2f%2fwww.sciencedirect.com%2fscience%2farticle%2fpii%2fS1574119216000444)

Machine Learning Techniques for Anomaly Detection: An Overview <https://pdfs.semanticscholar.org/0278/bbaf1db5df036f02393679d485260b1daeb7.pdf>