

Group Project Status Report for ‘Machine Learning Mavericks’

So far, our group, consisting of me (Bryan Sanchez), Joshua C. Meza, Ryan A. Schultz, and Preston D. Satterfield, has almost completed task 1, and has begun working on task 2. With respect to task 1, the logic for method 1 (frequency-based intensity) is done, using the ‘SciPy’ package in Python and performing a weighted non-parametric density estimation using a gaussian kernel, and we are attempting to find the best way to visualize the intensity estimations. Apart from just a simple histogram-like representation, we are also evaluating the intensity estimations over a grid of coordinates and using a color mesh to represent them. The logic for method 2 (energy-based) is inefficient, takes quite a long time to compute, and can be improved greatly. Since there is no convenient package, we are manually and iteratively computing the intensity estimation for each location in the 4-month sub-set of data by simply taking the average of the energy band range and multiplying it by its duration (which doesn’t yield the best results, so we may look to an alternative). As for its visualization, we are also applying the same techniques as we did to method 1.

With respect to task 2, since the hotspot detection is restricted to analysis based on method 1, we have begun developing our detection algorithm, as our method 1 is working well. The first idea for determining thresholds that we are attempting is taking the average minimum and maximum intensity from our results in task 1, creating a range using those, and finding ‘percentiles’ within that range to serve as reasonable thresholds. However, not much progress has been made on this so far. Given our progress with task 1 and 2, we have not yet done any kind of change analysis for task 3, and with respect to the full 8–10-page report, we have not yet started it, as we don’t feel like we have made a substantial amount of progress, and believe that we should at least have tasks 1 and 2 mostly done before starting. Barring any major setbacks or problems, we are still reasonably confident that we can finish the entire project before the November 10 deadline.