**CSCE 623: Machine Learning**

**Spring 2020**

**PROJECT PROPOSAL GRADING WORKSHEET**

Due Thursday, 14 May at 2359

Submit via Canvas

**(**This Project Proposal is worth 5 points toward your final grade**)**

|  |  |
| --- | --- |
| Course points earned | 5.0 |

**demoremark\_Project\_Proposal\_Grading\_Worksheet.docx**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Step | Step Evaluation Criteria | Avail points | Student performance | Student Score |
| 1 | Background: Student identified domain of study they are going to work in, and the specific problem they are going to address and why it is important - in sufficient detail | 1 | Achieved | 1 |
| 2 | Formal Task Description: Formal description of task using ML language – for example: classification, regression, supervised vs. unsupervised; bias, variance. | 1 | Achieved | 1 |
| 3 | Data Explanation: identified whether using existing data, generating new data.  Source of data example: datasets; simulations to generate the data; websites such as Kaggle or UCI data repository;  Data wrangling required?  Observations & feature count & description | 1 | Achieved | 1 |
| 4 | Assessment & Contribution:  if supervised; where getting truth information?  How will you evaluate performance?  How will your project support research? | 1 | Achieved | 1 |
| Q | Quality/Effort: Proposal follows formatting guidelines and length constraint.  Student appeared to have put sufficient thought and time into developing the proposal.  Proposal is well written and flows logically. Proposal is easy to understand. | 1 | Good | 1 |
| TOTAL |  | 5 |  | 5 |

**Questions and Suggestions:**

You have a good description of the project and thus earn full credit… however, a word of caution regarding AVAS data: Lt Marvin was very dubious about the veracity and usefulness of this spoofed data for Machine Learning… as was his advisor (Dr. Graham) and the thesis committee. While you can try a project which uses this data in CSCE 623, You should definitely have a conversation with Dr. Graham before proceeding to use this data for your thesis if you will be doing ML as part of your research. Be very careful in how you define and implement “spoofing” in any method of generating new data in order for the project not to just devolve to become a “disagreement detector” which can determine whether the INS is in too much disagreement with the GPS – such ML models are easy to build but don’t really estimate the performance of how the spoofer would work in real situations… especially if the GPS spoofing you choose to record is making the A/C appear to do unrealistically wildly different flight activities than what the INS reports. The trick is being able to use ML to detect a spoof which is so subtle that other methods would still attribute the disagreement between GPS and INS solely to INS drift. And do this without having too many false positives (the ML says there is a spoof when there isn’t a spoof) If you can achieve that, you will make a great contribution.