|  |  |
| --- | --- |
| **A Practical Application of Machine Learning-Based Classification Techniques to Proactively Identify Insider Threats**  Insider threats are on the rise. Current commercial software can monitor, log, and prevent access to designated files and directories. However, it remains difficult to predict and prevent unauthorized insider usage. Due to the gaps in research in the area, the focus of this study is to more accurately predict insider threats in a server environment.  The method utilizes the Amazon Machine Learning software. It will be trained on a dataset comprised of normal user situations, crafted mistakes, and malicious activity. The software will use the training dataset, to make predictions against similar datasets to verify accuracy. It will then be tested against a human actor to test predictions in a high-fidelity simulation. A testing dataset of 100 commands will be given to the model for batch prediction and will be evaluated for statistical accuracy.  The results of the test indicate the model isn't sufficiently accurate in detecting intrusion even when given special weighting. This is likely due to the lack of data in the malicious and mistake categories. As other researchers have indicated, larger datasets comprising of more points such as timestamps and keyboard information would help better track user behavior in an attempt to establish recognizable patterns by neural networks. This and other research indicate that more data is needed.  **Get the materials, source code, and references at:**  <https://github.com/skraelingjar/cyberml>  Joshua Bowen  Northern Arizona University - Prescott Valley  [jdb567@nau.edu](mailto:jdb567@nau.edu)  [http://joshuabowen.info](http://joshuabowen.info/) | Notes: |