School of Engineering and Computing Sciences

Comparison of Various Modeling Techniques for Driver Data

REU fellow:, Angela Burton[1], Faculty mentor: Dr. Jonathan Voris[2,]
Affiliation: 1. Vanderbilt University, 2. School of Engineering and Computing Sciences, NYIT
Email: 1. angela.j.burton@Vanderbilt.edu, 2. jvoris@nyit.edu
NYIT Research Experience for Undergraduates (REU)
May 29 – August 5, 2016

Summary

Cars usually employ token-based authentication. In some instances, however (ex. for car theft detection [5], insurance companies [4], car sharing services), knowing exactly who is driving is important. One way to achieve this is by analyzing the driver's behavior to ensure that it matches the usual behavior of the expected driver. A previous study collected driver data and developed a multiclass machine learning model that could accurately differentiate between drivers. We continued to analyze this data and developed a more accurate one-class model.

Behavioral Biometrics

- Behavioral biometrics authenticate the user continuously throughout a session (as opposed to one-time fingerprint or vocal recognition)
- Compares user behavior to an established profile

Data Collection

- Collected by previous study [1]
- Twelve features, including car position, car rotation, steering wheel rotation, brake and accelerator position
- OpenDS Driving Simulator
- Twelve participants
- Data collected every 10 ms, then processed into 10 second intervals

Research Tools

Weka
- GUI-based machine learning tool
- Used by the previous study [1]
- Difficult to produce graphs, missing some required functionality

MATLAB
- Used the Machine Learning and Statistics toolbox
- Includes a variety of modeling algorithms
- Outputs values necessary to produce a variety of graphs

Multiclass Support Vector Machine (SVM) Model

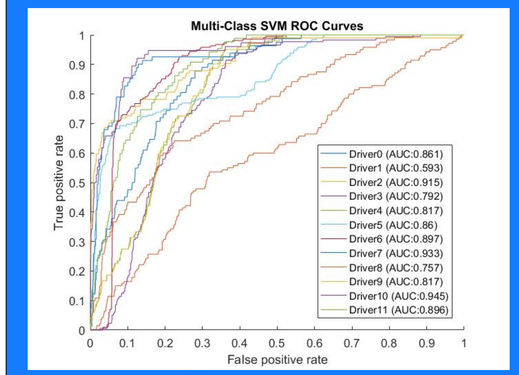

Figure 1. ROC curve for multiclass SVM model.

- Average area under the ROC curves: 0.840
- Average Equal Error Rate: 23.4%

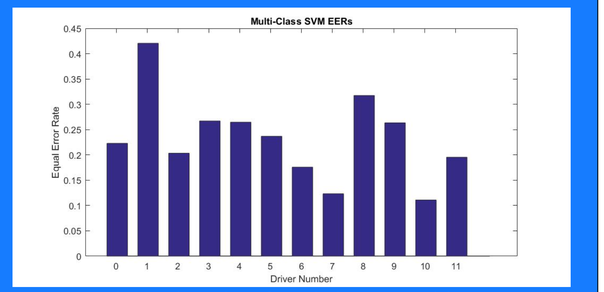

Figure 2. EER bar graph per driver for multiclass SVM model.

Multiclass Decision Tree Model

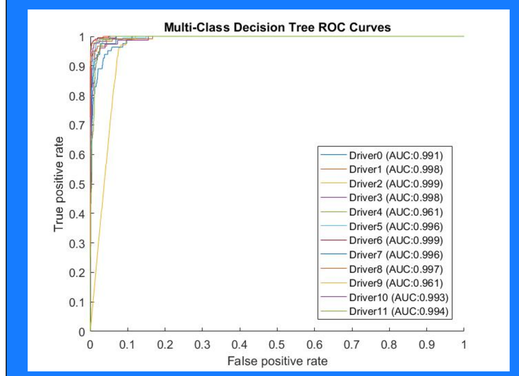

Figure 3. ROC curve for multiclass decision model.

- Average area under the ROC curves: 0.990
- Average Equal Error Rate: 4.68%

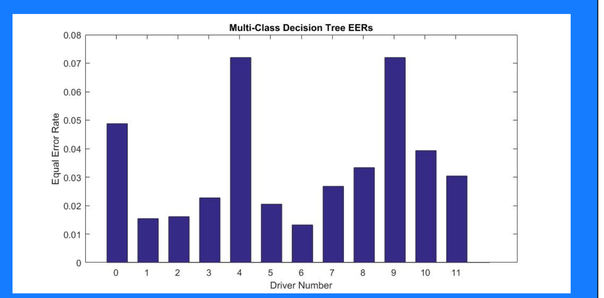

Figure 4. EER bar graph per driver for multiclass decision tree model.

One-Class SVM Model

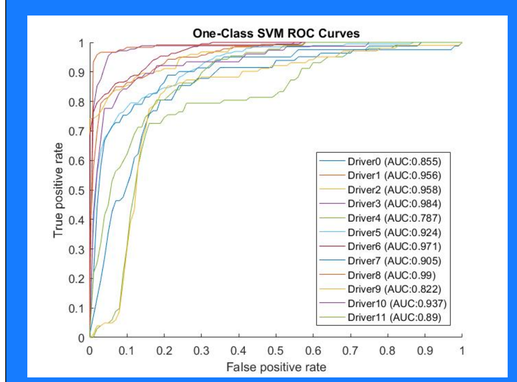

Figure 5. ROC curve for one class SVM model.

- Average area under the ROC curves: 0.915
- Average Equal Error Rate: 15.3%

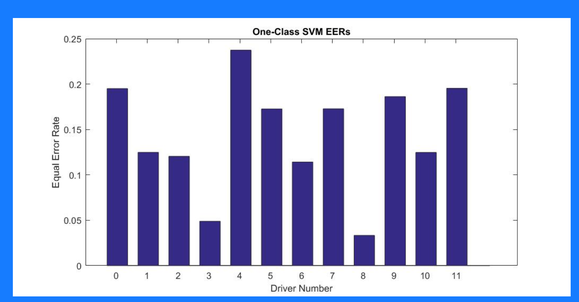

Figure 6. EER bar graph per driver for one class SVM model.

Methodology & Discussion

Multiclass SVM
- Same results as the previous study achieved [1]
- Produced better graphs of these same results

Multiclass Decision Tree
- Far better results than expected
- This is likely due to a peculiarity in the dataset

One-Class SVM (ocSVM)
- Most closely mimics the work that would be done by an in-car authentication system
- This system wouldn't have access to all driving patterns, only those of that particular car's drivers

Other
- Worked on, but did not complete, a Gaussian Mixture Model classifier

Future Work

- Additional experimental comparisons of modeling algorithms and parameters
- Continue to work on a Gaussian Mixture Model
- Continue to improve the ocSVM model

Acknowledgement

The project is funded by National Science Foundation Grant number CNS-1559652 and New York Institute of Technology.

References:

1. T. Parikh, S. Mascarenhas, J. Zhang, J. Voris, N. S. Artan and W. Li, "Modeling Driving Patterns for Enhanced Driver Authentication," *Under Review.*
2. M. Jensen, J. Wagner and K. Alexander, "Analysis of in-vehicle driver behavior data for improved safety," *Int. J. Vehicle Safety,* vol. 5, no. 3, pp. 197 - 212, 2011.
3. C. Karatas, L. Liu, H. Li, J. Liu, Y. Wang, S. Tan, J. Yang, Y. Chen, M. Gruteser and R. Martin, "Leveraging Wearables for Steering and DriverTracking," in *IEEE INFOCOM: The 35th Annual IEEE International Conference on Computer Communications, 2016* [Online]. Available: ResearchGate, https://www.researchgate.net.
4. D. Jergler, "Researchers question privacy of usage-based auto insurance," 2013, http://www.insurancejournal.com/news/national/2013/10/02/307073.htm
5. M. Salemi, "Authenticating users based on driving behavior," Ph.D. dissertation, Rutgers University-Graduate School-New Brunswick, 2015.