David Pagels

Aviation Data Mining

Key Points:

The question being addressed in this research is to figure out how to go about finding patterns in massive amounts of aviation flight recordings and incident data to prevent accidents from occurring in the future. This problem is being addressed by the use of data mining on flight recording data and aviation incident reports using several different data mining methods. These methods include Multiple Kernel Learning to find patterns in flight recording data and Natural language processing to classify text from incident reports.

The supporting sources of this paper that I've chosen address this problem from a couple of different angles. *Multiple kernel learning for heterogeneous anomaly detection* and *Detection of Precursors to Aviation Safety Incidents Due to Human Factors* both try to find anomalies in flight recording data. They do this by using Multiple kernel learning and miscellaneous anomaly detection methods, respectively. *Experiences in mining aviation safety data*, *Semi-Supervised cause identification from aviation safety reports*, and *Building the Prediction model from aviation incident data* all discuss the topic of mining information from aviation incident reports by using natural language processing for text classification. *Data Mining approaches for aircraft accident prediction* finds patterns in fatal aircraft accidents, but will likely be used as a supporting paper, as it is very background rich. By studying these two general methods of aviation data mining, we can analyze the effectiveness of each in relation to each other.

Data mining in aviation is a young, but growing, field. It has been used successfully to find sequences of events or anomalies that have a heightened chance to end in an accident. As data mining techniques are improved, it allows for more patterns to be found, becoming even more effective.

In order to explain some of the data mining methods used, the reader will need to be informed of concepts such as kernels and support vector machines.

Outline:

Data Mining in Aviation

1.Introduction

Introduce the concept of data mining in aviation and a rough description of the direction of the paper

2.Background

This is where I will describe concepts such as kernels, natural language processing, and support vector machines.

- 2.1 Concepts
- 2.2 Kernels/ Support Vector machines
- 2.3 Natural Language Processing
- 2.X Other needed background info

3. Methods

- 3.1 Data Mining Flight Recording Data
  - Multiple Kernel Learning … will be the main source for this section (supported by support vector machine background)
- 3.2 Data Mining Aviation Incident Reports
  - Semi-Supervised Cause Identification… will be the main source for this section (supported by Natural language processing background)

4.Results

Discussion of what has come from these two methods of aviation data mining

5.Conclusion

Discussion of the direction and effectiveness of the two fields with some comparison between the two.

6. Acknowledgements

7.References

Santanu Das, Bryan L Matthews, Ashok N Srivastava, and Nikunj C Oza. 2010. Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 47-56.

Igor Melnyk, Pranjul Yadav, Michael Steinbach, Jaideep Srivastava, Vipin Kumar, and Arindam Banerjee. 2013. Detection of Precursors to Aviation Safety Incidents Due to Human Factors. In Data Mining Workshops (ICDMW), 2013 IEEE 13th International Conference on. IEEE, 407-412.

AB Arockia Christopher and S Appavu. 2013. Data mining approaches for aircraft accidents prediction: An empirical study on Turkey airline. In Emerging Trends in Computing, Communication and Nanotechnology (ICE-CCN), 2013 International Conference on. IEEE, 739-745.

Alexandra Lukacova, Frantisek Babic, and Jan Paralic. 2014. Building the prediction model from the aviation incident data. In Applied Machine Intelligence and Informatics (SAMI), 2014 IEEE 12th International Symposium on. IEEE, 365-369.

Zohreh Nazeri, Eric Bloedorn, and Paul Ostwald. 2001. Experiences in mining aviation safety data. In ACM SIGMOD Record, Vol. 30. ACM, 562-566.

Isaac Persing and Vincent Ng. 2009. Semi-supervised cause identification from aviation safety reports. In Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2-Volume 2. Association for Computational Linguistics, 843-851.