Analysis of US Airline Stocks Performance Using Latent Dirichlet Allocation (LDA)

Graduate Researcher : Amina Issoufou Anaroua, Master of Science Aviation Finance
Faculty Mentor : Dr. Hari Adhikari

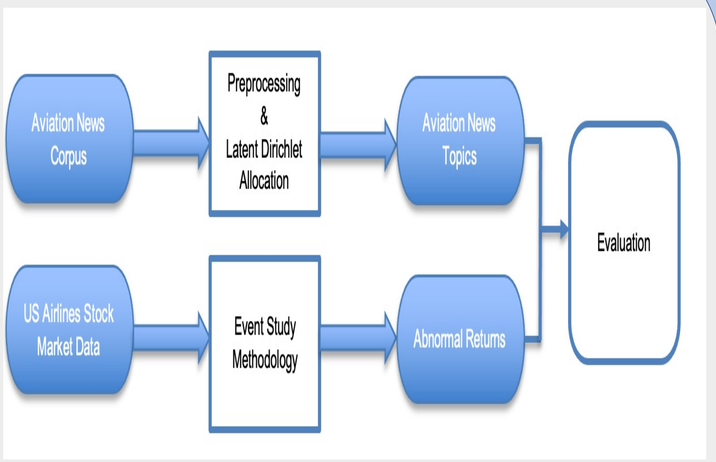
Abstract

Our data account for aviation news of airlines, airports, regulations, safety, accidents, manufacturers, MRO, incidents, aviation training, general aviation and others from Aviation Voice. First, we mine the aviation related data through text mining and topic modeling. Second, we employ the LDA model approach to help us identify and capture the extent of certain topics mentioned in aviation voice news releases. Finally, we use Event Study Methodology (ESM)to investigate stock price reactions to news announcements. Eventually, we succeed in extracting 10 topics.

Introduction

- Use aviation digitized information to analyze the performance of US airline stocks
- Qualitative information is contained in news announcements
- Data account for aviation news of airlines, airports, regulations, safety, accidents, manufacturers, MRO, incidents, aviation training, general aviation and others from Aviation Voice
- Use a natural language processing, Latent Dirichlet Allocation (LDA) to investigate and search for patterns that can explain the movement of US airline stock

Methodology

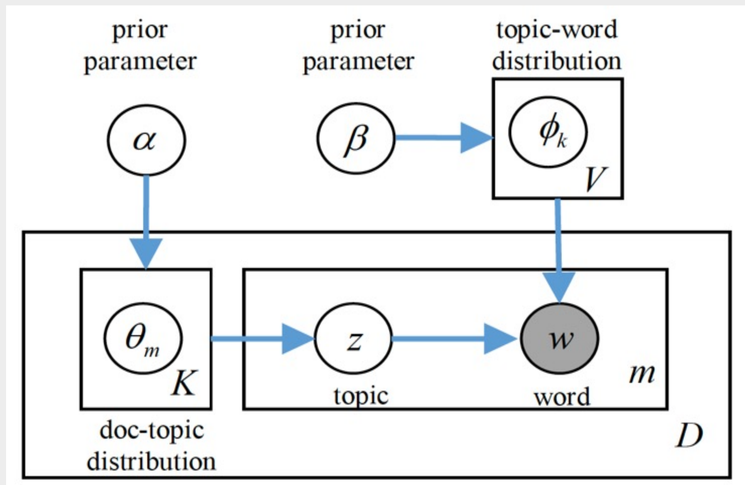


Mining Process

- Text Mining
  Data cleaning , Tokenization, Stemming
  Creation of document-word matrix
- Topic Modelling

Latent Dirichlet Allocation

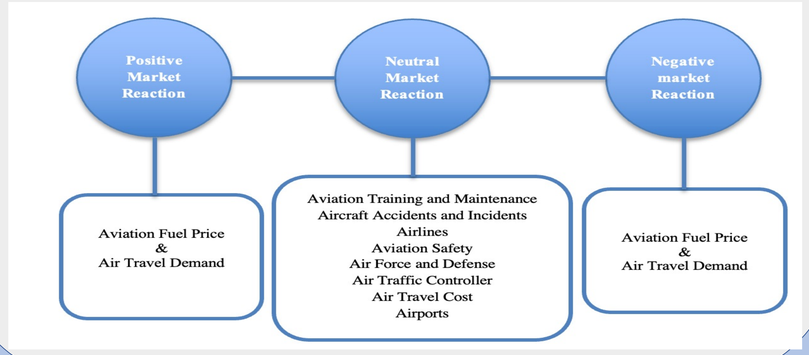
- For each document, a distribution over topics is selected randomly
- Next , the process continues for each word in the document



Findings/Data

2016 to 2020, 1716 news articles from Aviation Voice

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Topics
Topic 0	aircraft	flight	engine	test	say	aviation	design	technology	program	fuel	Aviation Fuel Price
Topic 1	pilot	training	aviation	maintenance	airline	say	program	need	flight	service	Aviation Training and Maintenance
Topic 2	flight	aircraft	plane	say	crew	passenger	report	crash	airport	engine	Aircraft Accidents and Incidents
Topic 3	aircraft	boeing	max	order	airbus	airline	boee	delivery	say	airplane	Airlines
Topic 4	aviation	faa	safety	drone	say	use	issue	pilot	process	datum	Aviation Safety
Topic 5	air	force	fighter	defense	lockheed	jet	state	mission	aircraft	japan	Air Force and Defense
Topic 6	say	engine	whitney	lufthansa	traffic	controller	air	house	fee	pratt	Air Traffic Controller
Topic 7	jet	business	company	charter	cost	travel	ita	price	hour	plane	Air Travel Cost
Topic 8	year	air	market	growth	demand	increase	airline	passenger	business	grow	Air Travel Demand
Topic 9	airline	flight	airport	passenger	carrier	service	fly	route	delta	air	Airports



Conclusion

- Content of news announcements conveys information that are reflected in the stock market prices
- Not all the news that have impact or transmit important message on the stock prices
- Tagging of news articles categorizes articles superficially and is not sufficient

Future Research

- To get more comprehensive results, electronic articles from other aviation sources can also be included. Additional research can be conducted to extend the period of study and verify the robustness of our approach.
- Some news articles have strong relationship among them, but LDA model suppose that documents are independent from one another. When inferring the topics, the position of individual words is disregarded. Different methods of topics extraction should be studied to perform a comparison with the LDA.

References

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. The Journal of Machine Learning Research, 3, 993-1022.

Feuerriegel, S., Ratku, A., & Neumann, D. (2016). Analysis of how underlying topics in financial news affect stock prices using latent dirichlet allocation. In 2016 49th Hawaii International Conference on System Sciences (HICSS), 1072-1081.

For more Information

Contact : anarouaa@my.erau.edu