



Graduate Researcher : Amina Issoufou Anaroua, Master of Science Aviation Finance

Faculty Mentor : Dr. Hari Adhikari

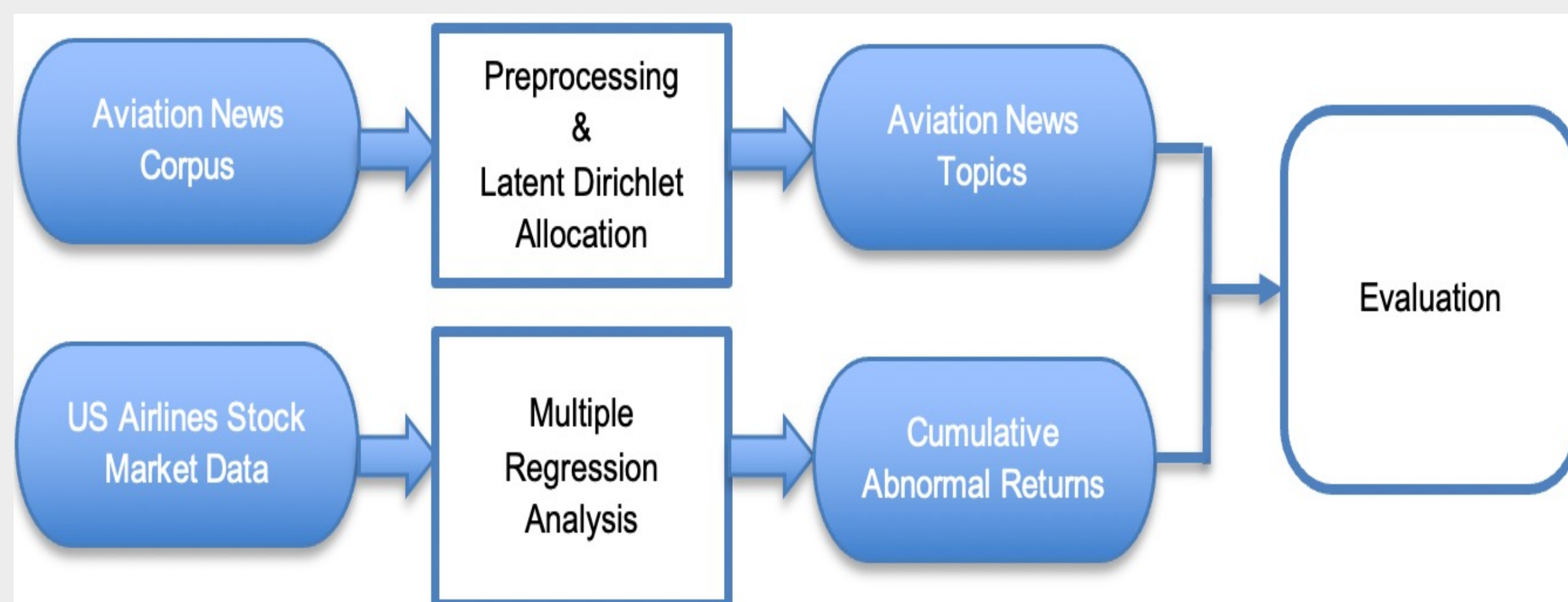
Abstract

We use Latent Dirichlet Allocation (LDA) to investigate and search for patterns that can explain the movement of US airline stock. 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 multiple regression models to investigate stock price reactions to news announcements. Eventually, we succeed in extracting 10 topics. As hypothesized, the impact of those topics varies greatly from topic to topic. Some of the topics have effect on the US Airline stocks in the short and long terms moving average while other topics have only effect on the medium-term run.

Introduction

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

Methodology



Mining Process

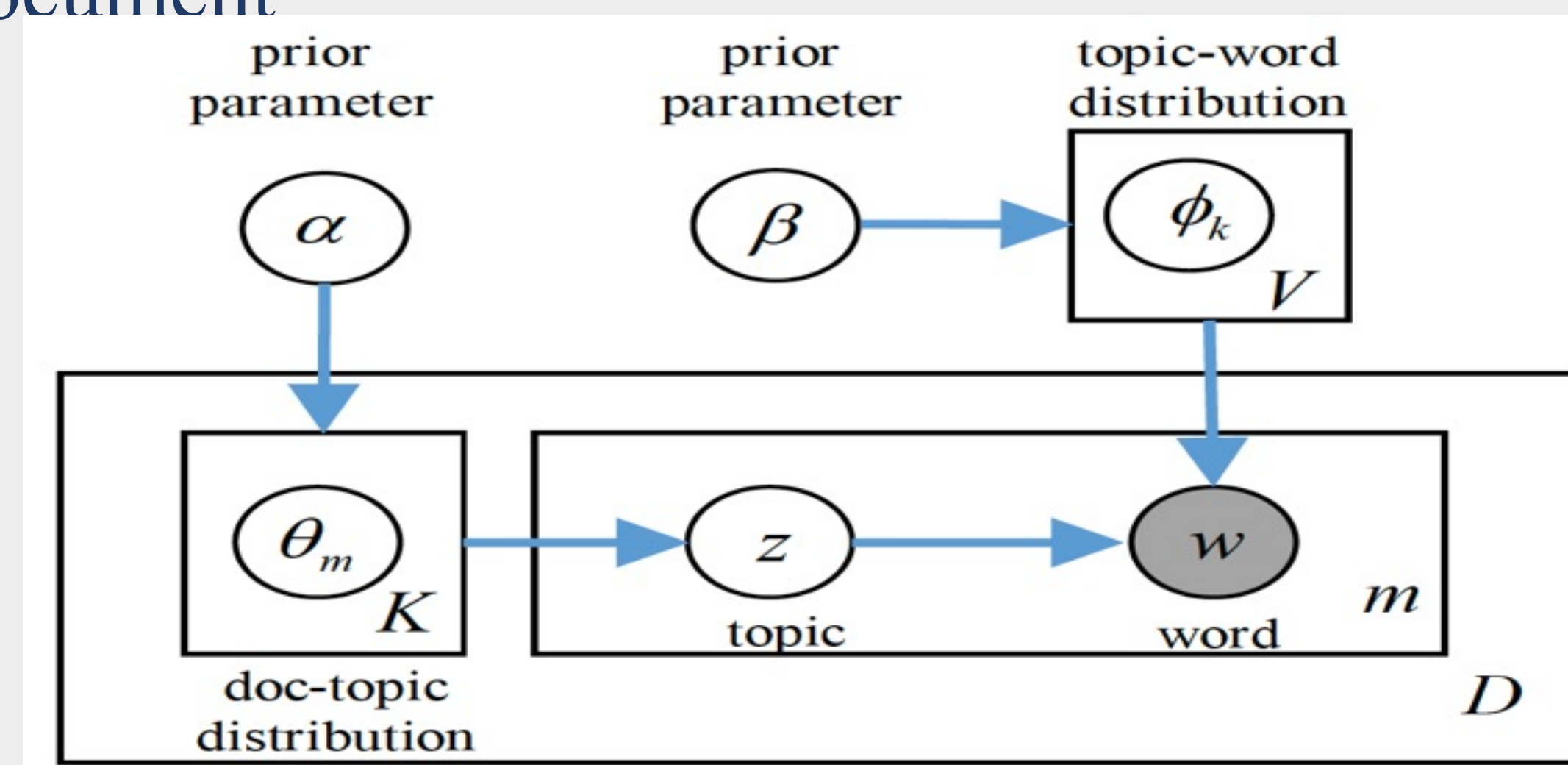
- Text Mining
Data cleaning , Tokenization, Stemming
Creation of document-word matrix
- Topic Modelling

Latent Dirichlet Allocation

Multiple Regression Analysis

Latent Dirichlet Allocation

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



Multiple Regression Analysis

$$\text{Log (CAR}_j) = \alpha + \sum_{i=0}^j (\beta_{i,j} \text{Topic}_{i,j}) + u_{i,j}$$

- β, α , unknown parameters
- u , error terms
- i , from 0 to 9 for the 10 news topics
- j , for 3, 5, 7, 15, and 21-days rolling average of abnormal returns
- Topic, news topics
- CAR, cumulative abnormal returns

Conclusion

- Content of news announcements conveys information that are reflected in the stock market prices
- News transmit important message differently on the stock prices over time
- Tagging of news articles categorizes articles superficially and is not sufficient

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.

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	boe	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

	Topic0	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7	Topic8	Topic9
CAR ₀₃	+	+								
CAR ₀₅		+					-			
CAR ₀₇			-	-			-			
CAR ₁₅	+							+		
CAR ₂₁	+	+		+	+		+		+	+

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.
- When inferring the topics, the LDA model ignores the position of individual words. Different methods of topics extraction should be studied to perform a comparison with the LDA.

For more Information

Contact : anarouaa@my.erau.edu