

The Hong Kong University of Science and Technology

Department of Mathematics

Mathematics Colloquium

Predicting Returns with Text Data

Prof. Dacheng Xiu University of Chicago

<u>Abstract</u>

We introduce a new text-mining methodology that extracts sentiment information from news articles to predict asset returns. Unlike more common sentiment scores used for stock return prediction (e.g., those sold by commercial vendors or built with dictionary-based methods), our supervised learning framework constructs a sentiment score that is specifically adapted to the problem of return prediction. Our method proceeds in three steps: 1) isolating a list of sentiment terms via predictive screening, 2) assigning sentiment weights to these words via topic modeling, and 3) aggregating terms into an article-level sentiment score via penalized likelihood. We derive theoretical guarantees on the accuracy of estimates from our model with minimal assumptions. In our empirical analysis, we text-mine one of the most actively monitored streams of news articles in the financial system—the Dow Jones Newswires—and show that our supervised sentiment model excels at extracting return-predictive signals in this context. First, the model selects a list of positive and negative sentiment words that is clearly interpretable and intuitive. Moreover, a simple trading strategy demonstrates that our news sentiment scores (by a large margin).

Date: Friday, 22 March 2019 Time: 3:00p.m. - 4:00p.m. Venue: Lecture Theater F, Academic Building, 1/F (near Lifts 25 - 26), HKUST

All are welcome!