FRIDAYS:
A FINANCIAL RISK INFORMATION DETECTING AND ANALYZING SYSTEM

Chi-Han Du, Yi-Shyuan Chiang, Kun-Che Tsai, Liang-Chih Liu, Ming-Feng Tsai, Chuan-Ju Wang

CLIP&CFDA Lab, Research Center for Information Technology Innovation, Academia Sinica, Taiwan

Introduction

FRIDAYS, Financial Risk Information Detecting and Analyzing System, is a system integrating multiple NLP models trained on SCE's 10k financial reports that aims to assist professionals to digest financial reports in a more efficient manner. With the NLP results, the system can classify the tones of individual words, those on multi-words level, and the overall risk of any selected sentence.

Two main features, company overview and risk assessment, are present. The functionality of the former is set to provide users with holistic views and the general conditions of the selected companies; the later will focus on how to best present the NLP results for users to understand the financial reports.
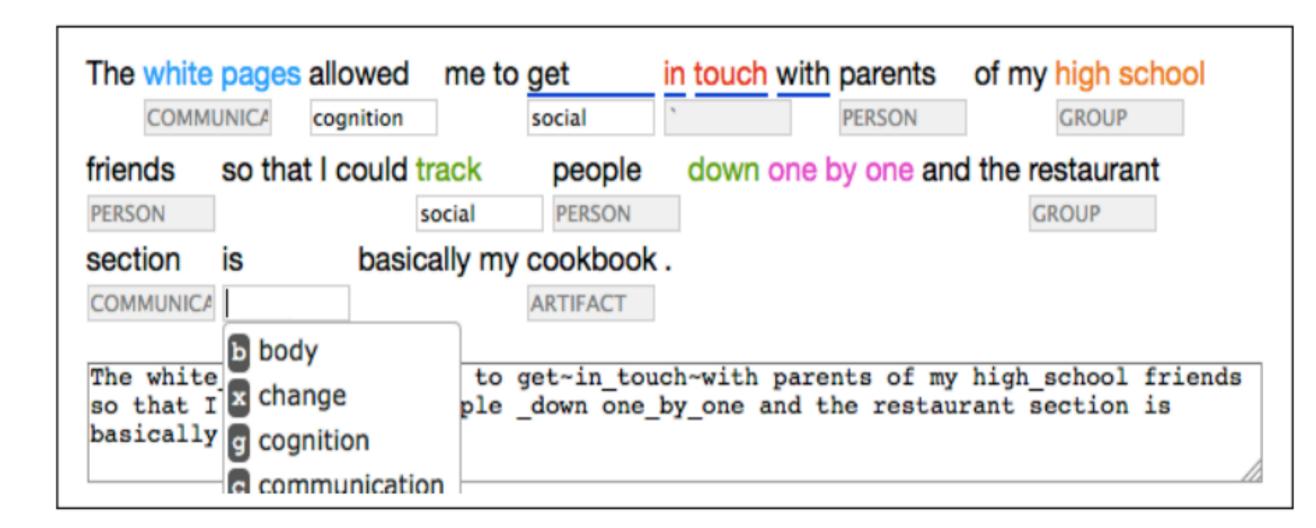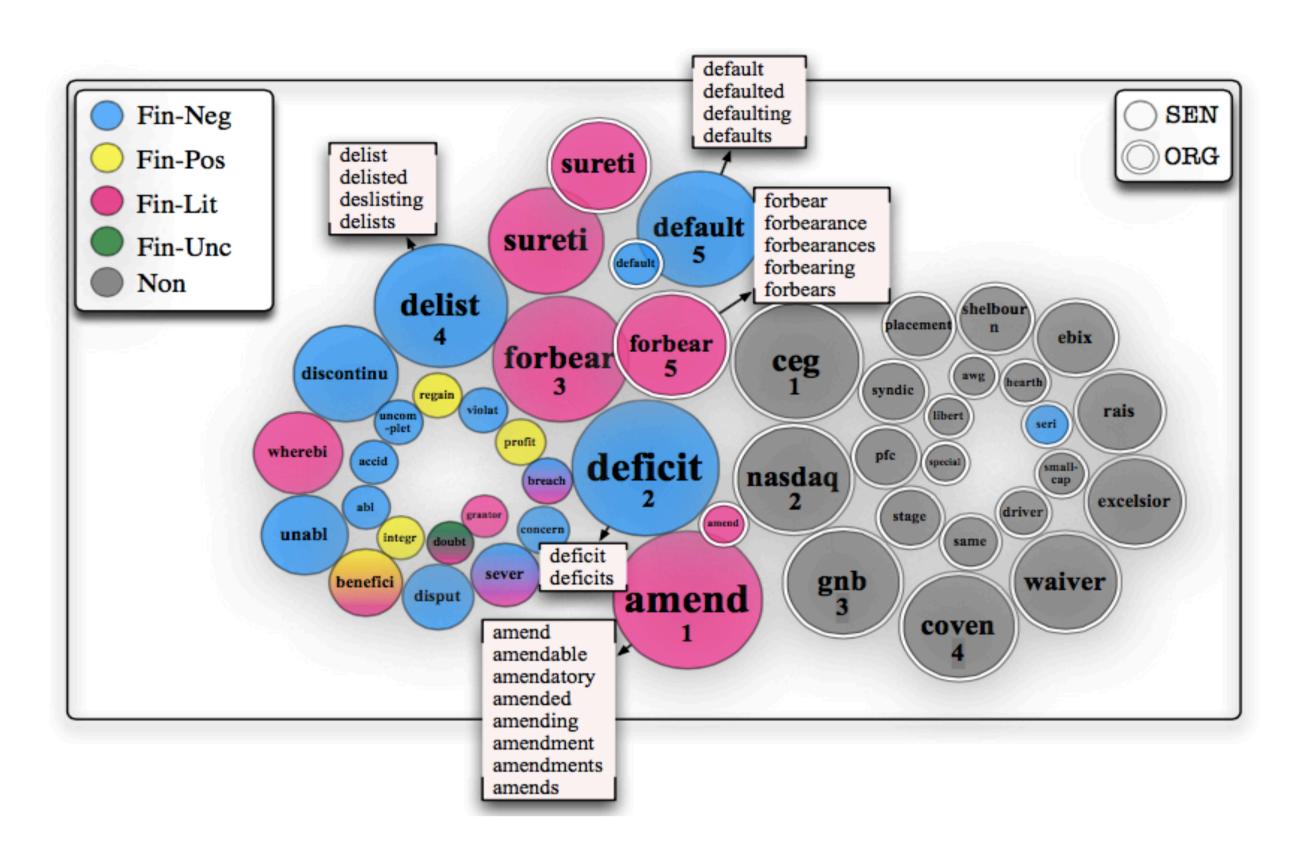
Dataset

Form 10-K dataset
Item 1 - Business
Item 2 - Properties
Item 3 - Legal Proceedings
Item 4 - Mine Safety Disclosures
Item 5 - Market
Item 6 - Consolidated Financial Data
Item 7 - Management's Discussion and Analysis of Financial Condition and Results of Operations
Item 8 - Financial Statements
Item 9 - Changes in and Disagreements

With Accountants on Accounting and Financial Disclosure

We analyze 40,708 annual SEC-mandated financial reports on Form 10-K from year 1996 to 2013. Besides, Our financial sentiment analysis is based on sentiment information of words from previous research including Loughran'11 and Wang'13 (above figure).
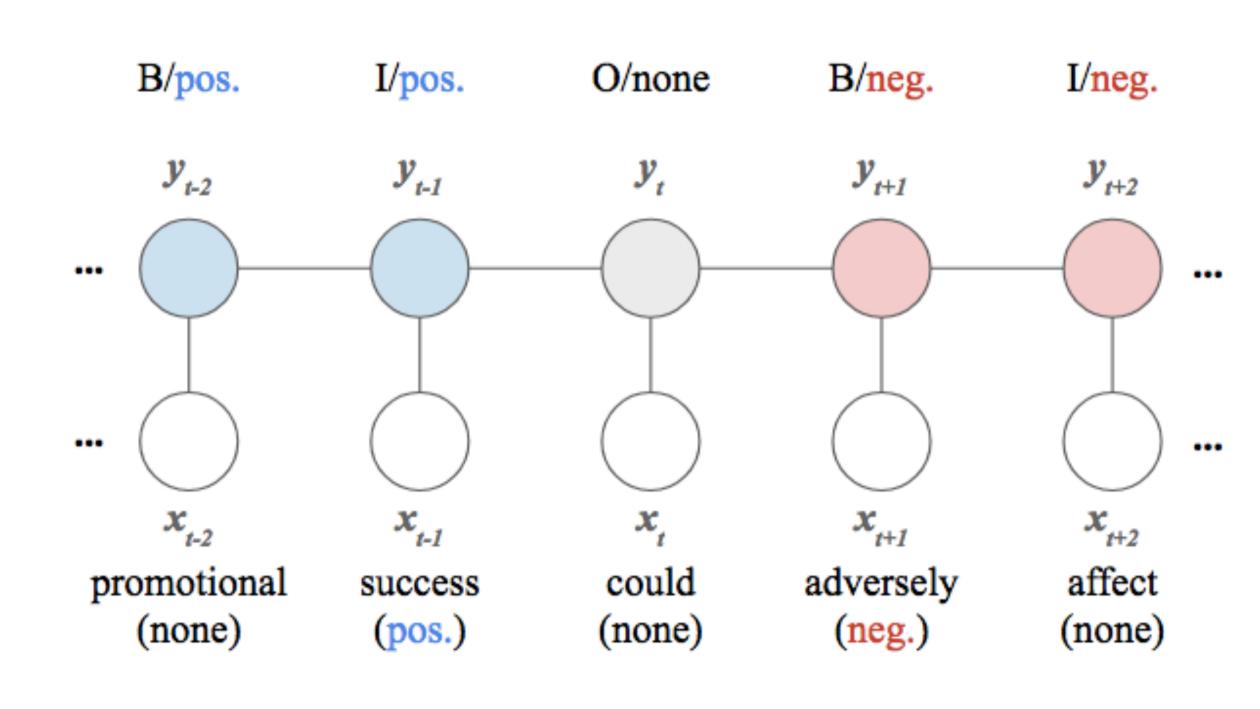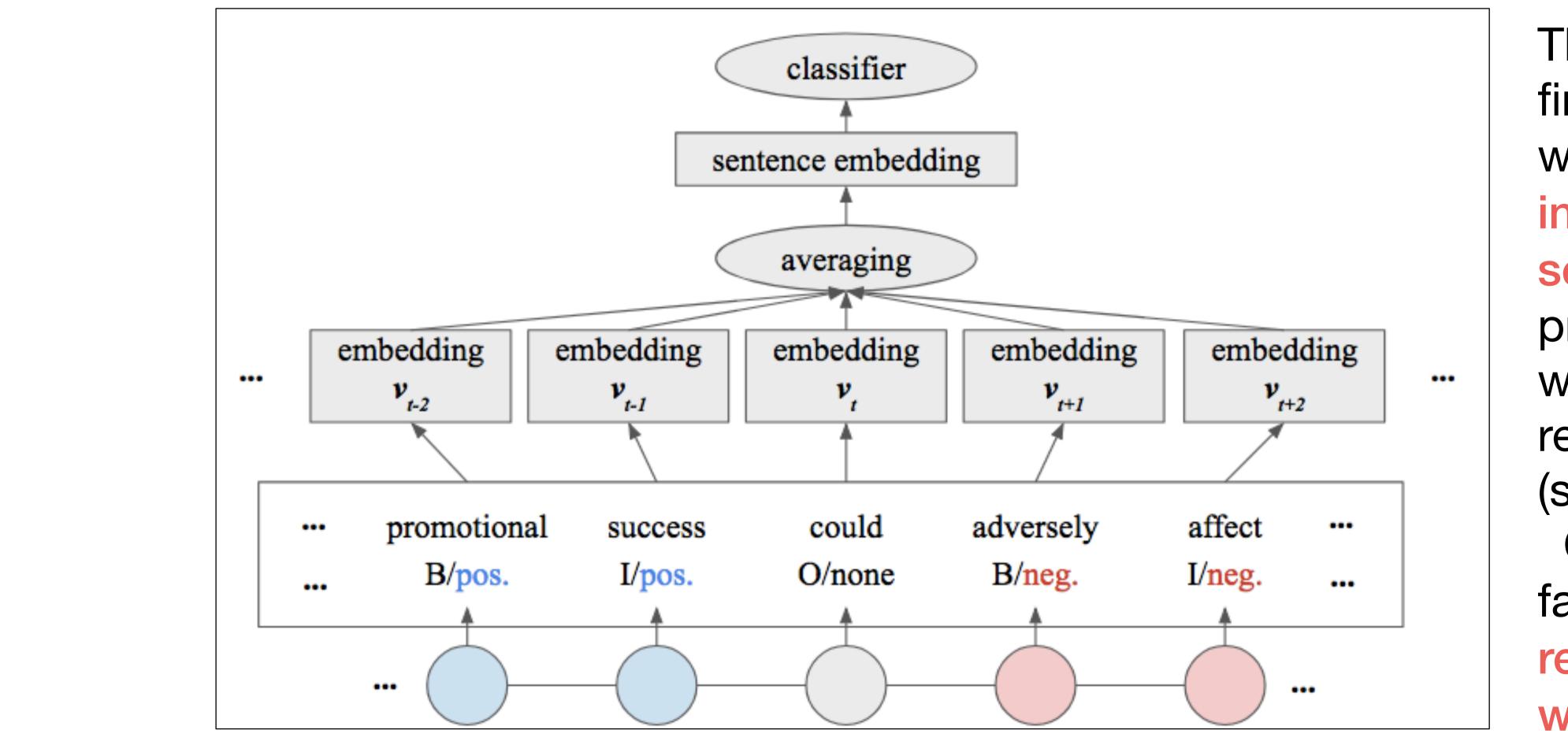
MWE sentiment detect model

Following the annotation procedure mentioned by Schneider'15 (below figure), we developed our own version of adoption (on financial texts) and purpose (for 4 sentiment classes) for MWE sentiment annotation.

In order to compete with SOTA MWE detection models using DL architecture, our lab colleagues are developing an advanced CRF model to predict both range and sentiment of MWE.

Financial risk classifier

The core module of FRIDAYS is the financial risk classifier model, which reads the word sentiment information together with MWE sentiment information, and predicts a financial sentence is whether high risk (larger stock return volatility) or low/neutral risk (smaller volatility).

Currently we use model trained in fastText, but it can be easily replaced by other DL models within our system.
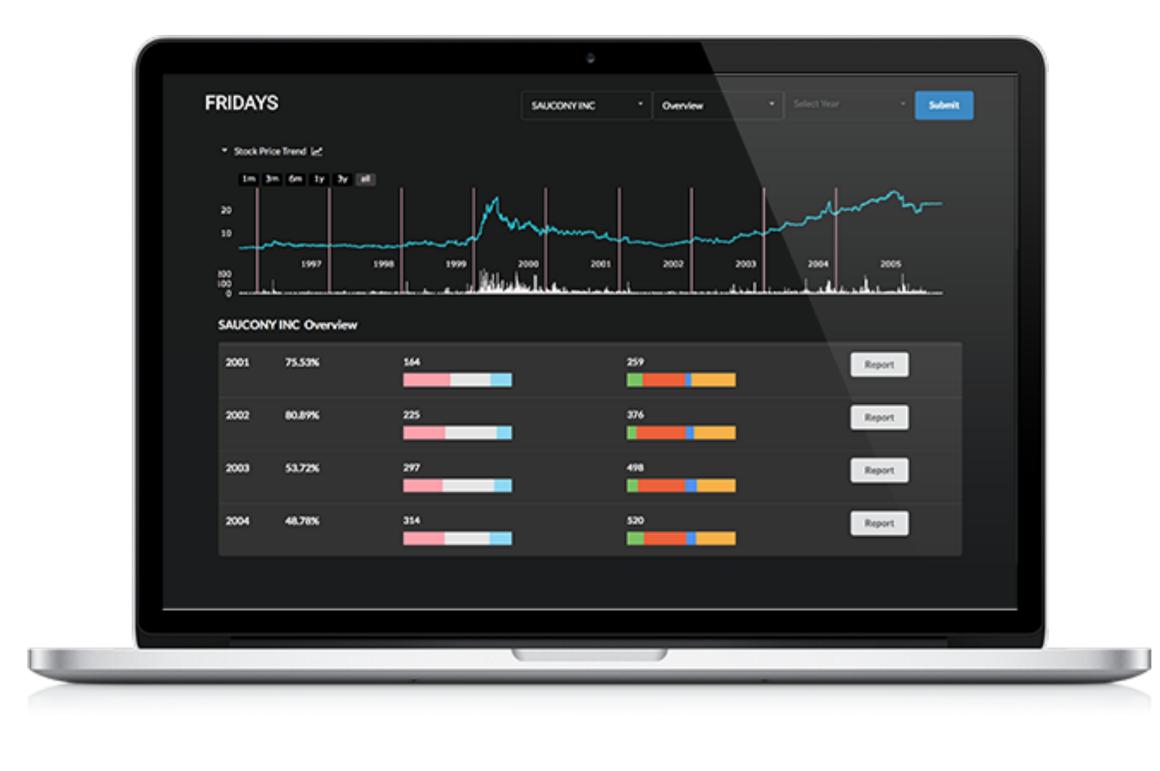
Use Case: Saucony Inc.

Imagine you are an investor interested in the company called Saucony Inc., and want to know whether you should go for its stock. It is impossible to get an overview from 9 annual reports in 5 minutes. You wonder, "how could I quickly digest convoluted sentence like this, and 300 more of them within hours?"
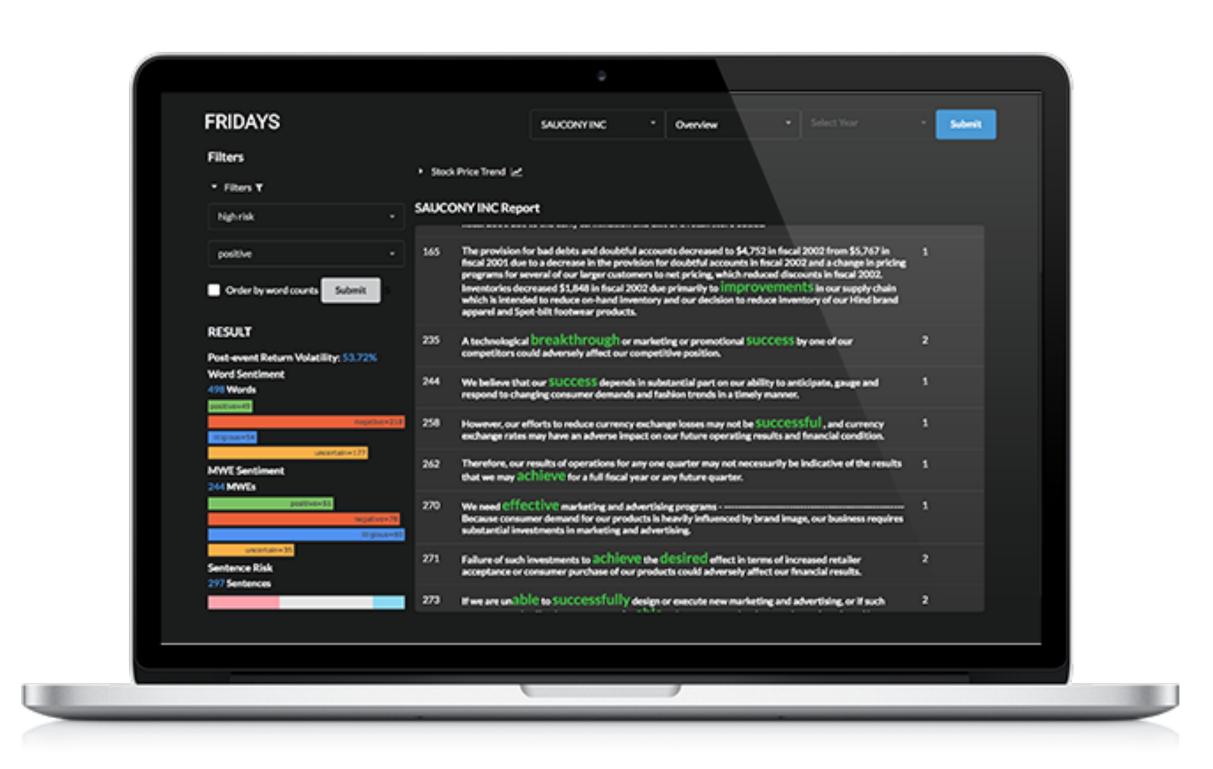
"A technological breakthrough or marketing or promotional success by one of our competitors could adversely affect our competitive position."
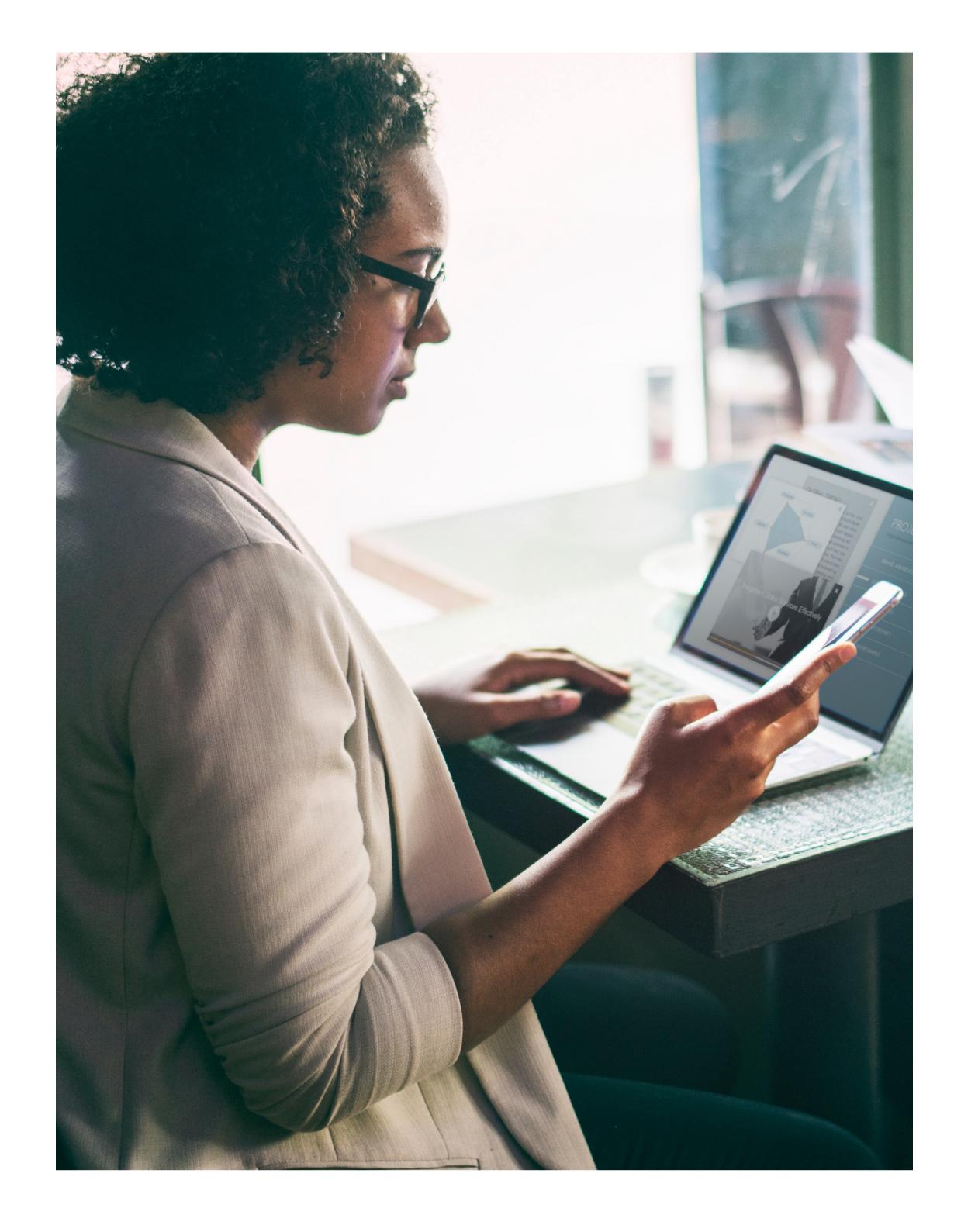
With FRIDAYS you can immediately know that its performance was more pessimistic after year 1999 since there exist more sentences expressing high risk; and even which sentence is vital by a single click!

You can even further investigate cases that was filled with promising terms but indeed risky! And, If you can rely on the system before invest your money, imagine how it could help your AI research on financial text when you are wondering whether your model "makes sense"?

https://cfda.csie.org/FRIDAYS/

User Interfaces

COMPANY OVERVIEW
The overview consists of every available financial report from the SCE dataset, accompanied with the post volatility of each year, the NLP analysis results and the direct access to the reports.

RISK ASSESSMENT
Risk assessment focuses on further report investigation, the NLP prediction results are rendered into charts as showed in the left column, while the corpus is on the right. All chart is interactive, and once a result type is selected, the corresponding text will be highlighted in the report.