BANYAN’S VISION

Banyan strives to change a century old industry by building a data-driven underwriting process

PROBLEM STATEMENT

Develop a model to predict claims and tools to help the vision

1. Is the industry looking at the right elements for prediction?
2. Can we predict claims with a small number of factors?
3. What are the factors that matter?

FACTOR CONSTRUCTION

Using academic research and Banyan’s experience we formulated predictive hypothesis for claims

Size comes with responsibilities
Once a liar, always a liar
Some owners have more litigations

DATA EXTRACTION PIPELINE

Developed automated data extraction pipelines from several sources: CapIQ, ISS.
Created a methodology to merge companies with securities, clean, enrich the data.

What is the market capitalization?
Was the company sued before?
Who is the owner?

MODELLING

HIGH DIMENSIONALITY
+8000 US companies
+100 Features
10 Years of data per company per feature

MODELS
CART
XGBOOST
RANDOM FORESTS

CLASS IMBALANCE
SMOTE

GLOBALISATION
3000 more companies for Canada

RESULTS

+16% Increase in AUC our best model versus the baseline

FEATURE ENGINEERING

Engineered transformations and variations of factor measures to construct predictive features.

Calculate the growth of market capitalization.
Extract the result of the litigation.
Capture changes in owners

IMPACT

+3 h saved per submission
+$9,000,000 in premium per person

The insurance industry is not looking at all the relevant factors – rigid view on market cap and sector

For large companies: size, short interest, volatility, employee growth, institutional ownership matter

For small companies: size, healthcare sector, financial sector, quality of earnings matter

WEB APP

Search any public company in the US by its ticker