**TITLE**

Bridging The Gap Between Machine Learning, Fundamental Analysis and Crowdsourcing

**ABSTRACT**

…….

**INTRODUCTION**

* Problem Statement: The investment management industry is facing numerous challenges and pressures. Data availability and complexity is growing, the stock market is becoming increasingly efficient, management fees are being compressed and competition continues to rise in the form of active and passive investment products. On the one side of the industry are investment managers that are failing to embrace the technological innovation in the form of AI and data science tools. And on the other side are quantitative managers that are scouring the data universe for new sources of alpha generation. Rarely do fundamental and quantitative managers bridge the gap between their expertise in a collaborative and effective manner as both are entrenched in their respective camps for numerous reasons thus we believe a chasm remains. In order to survive and thrive in this dynamic and challenging landscape, investment firms need tools that are: systematic, understandable, adaptable, complementary, efficient, scalable, multidimensional, innovative and effective. We are proposing the bridging of the gap between fundamental and quantitative insights and tools by bringing together domain knowledge and crowdsourced information with quantitative tools primarily in the form of machine learning and neural networks.

**LITERATURE REVIEW**

* Hypothesis: The investment management industry will be pressured to embrace the use of data science tools in order to survive and thrive. Humans and computers have their strengths and weaknesses and will be best served to leverage the strengths of both in order to best evolve and compete effectively. We believe the combination of domain knowledge, crowdsourcing of information and data science tools are a recipe for success. Our research review is focused on the following topics:
  + Quantitative investing overview
    - BofA Quant Primer
  + Current usage of machine learning in the investment industry
    - 10 reasons ML Funds Fail
    - ERP ML
    - AI Pioneers in Investment Management
  + Academic research on machine learning in the investment industry
    - A Machine Learning View on Momentum and Reversal Trading
    - A Backtesting Protocol in the Era of Machine Learning
    - Fundamental Analysis via Machine Learning
    - Machine Learning-Based Financial Statement Analysis
    - Machine Learning for Stock Selection
    - Predicting Profitability Using Machine Learning
  + Academic research regarding analyst revisions
    - The Information Content of Financial Analysts Forecasts of Earnings
    - Herding
    - Analyst’s stock views and revision actions
    - Financial Analysts’ Forecasts of Earnings
    - Herding among security analysts
    - Seemingly inconsistent analyst revisions
    - The behavioral basis of sell-side analysts herding
  + Academic research regarding momentum phenomena in the stock market
    - Momentum Strategies
    - The role of analyst forecasts in the momentum effect

**METHODS AND DATA**

* The focus of our analysis will be to develop predictive models that focus on the following aspects of a given investment:
  + Valuation
  + EPS
  + Revenues
  + Price
* Categories on model predictions to be made for a given metric:
  + Level
  + Trend
  + Relative level
  + Relative trend
  + Relative classification
* Initial methods we may use based on preliminary analysis
  + Machine learning algorithms
  + Neural network algorithms
* Data source
  + FactSet database
    - Through the SMU Business Library we gained access to FactSet’s database via an excel add-in.
    - We downloaded specific data on ~660 US large cap stocks going back to 12/31/09.
    - We then created numerous custom metrics based on the downloaded data.
    - Categories of our financial data include:
      * Valuation
      * Analyst revisions
      * Analyst EPS estimates
      * Analyst revenue estimates
      * Price
* Upon developing models we will then determine recommended usage and application of such tools as well as visualization tools for users of our models.
* Then we will create a scenario analysis whereby we apply our models and tools to the years 2020 and 2021, as we “take on COVID”

**RESULTS**

* Summarize model building and visualization efforts
* Summarize scenario analysis
* Accept, reject or modify hypothesis based on the results
* Clarify study scope

**DISCUSSION**

* Discuss modeling efficacy
* What unexpected insights did we garner
* Discuss takeaways – wins and losses
* What challenges did we face
* Make recommendations as to how our research can be applied in the industry
* Discuss aspects where others can go deeper with our research
  + Valuation insights
    - Having more valuation information vs peers using DS to understand best way to determine relative value
  + Insider transactions
  + More technical variables
  + More macro variables that directly or indirectly affect companies

**CONCLUSION**

* Overall summary of how this research is useful
* Revisit opening statements regarding the state of the industry and addressing of needs

**REFERENCES**

* ……