Theme Retrieval on Natural Language

Augmenting investment research with natural language processing for thematic funds

Morningstar Quantitative Research
22 April 2022
Version 1.0

Contents

Prashant Srivastava
Lead Data Scientist
prashant.srivastava@morningstar.com

Kenneth Lamont
Senior Quant, Manager Research
kenneth.lamont@morningstar.com

Monika Dutt
Director of Manager Research
monika.dutt@morningstar.com

Sabeeh Ashhar
Director of Quantitative Research
sabeeh.ashhar@morningstar.com

Executive Summary

In the current digital age, mutual fund and ETF data is vastly unstructured and accessible in disparate formats from a variety of sources, such as fund prospectuses, asset managers' websites, investment platforms, news outlets, data vendors, social media posts, and so on. As data on funds continues to expand, it is becoming more challenging for investors to process this growing body of information.

As the number of thematic funds has multiplied in recent years, so has the unstructured data connected with those funds. To help investors make better-informed decisions, Morningstar developed a global taxonomy for thematic funds in 2019 based on intentionality. To identify intentionality, we have relied on a combination of fund names (a strong indicator of intentionality) and information gleaned from prospectuses, marketing materials, index methodologies (in the case of index funds), and data points available within Morningstar Direct, such as investment objectives, where possible.[1]

Our team applies emerging technologies, such as the rising volume of fund data and the computer power of machines, to augment our research outcomes. In 2021, we enhanced our dataset by using natural language-processing technology to scan Morningstar's comprehensive global fund database. As a result, we have been able to construct an even more complete picture of the global thematic fund market than was previously possible.

In this paper we provide an overview of TRON, or Theme Retrieval on Natural Language, model. Our findings indicate that this technique significantly improves the efficiency of our Thematic Tagging process using Natural Language Processing, or NLP, techniques.

Introduction

More unstructured data exists today than ever before. According to the International Data Corporation, we create more data globally every two days now than we did from the beginning of civilization through to 2003.[2]

As the mountain of information continues to grow exponentially, it is becoming increasingly difficult for investment analysts to parse financial data efficiently, quickly, and meaningfully. A survey conducted by Deloitte found that analysts "spend roughly two-thirds of their time, on average, collecting and

1 Morningstar Global Thematic Funds Landscape 2022. Morningstar Manager Research. March 2022.
https://www.morningstar.com/lp/global-thematic-fund-landscape

2 Goldstein, I., Spatt, C.S., and Ye, M. "Big Data in Finance" (March 22, 2021).
SSRN: https://ssrn.com/abstract=3809447 or http://dx.doi.org/10.2139/ssrn.3809447

understanding data before knowing whether the information is material."[3] With limited time, researchers may not be able to identify all the relevant information that they need to make investment decisions.

The regulatory environment has also been a headwind for research departments in financial institutions. MiFID II, the European Union's legislative framework for financial markets, requires buy-side firms to pay for research services. Prior to MiFID II, sell-side brokers were allowed to bundle their research with other services. But this is no longer allowed. As a result of these structural and regulatory changes, there is a growing need for financial-services companies like Morningstar to provide you, our clients, with analytical tools that can process data and make inferences at scale.

Artificial Intelligence, or AI, and cloud computing have revolutionized the asset-management industry. AI algorithms have made it possible to sift through billions of financial data points and generate insight quickly and with a reasonable level of accuracy. While these monstrous models sometimes need significant computing power, they can now be easily scaled with cloud computing. This has made it possible for the investment-management industry to implement these technologies at a rapid pace.

In recent years, research in the NLP arena has increased significantly. This computer science technique can make meaningful inferences using written text and spoken language as inputs. The financial-services sector mainly uses NLP techniques to extract information or sentiment analysis from textual data, such as documents, social media posts, or news to better predict financial market movements.[4] In this report, we explain how we apply NLP techniques to yet another use case: thematic fund classification.

Investor interest in thematic funds has increased dramatically in recent years, particularly since the beginning of the coronavirus pandemic. In the trailing two years to the end of 2021, assets under management in these funds have grown nearly threefold to USD $806 billion worldwide.[1] As such, there is an increased need to identify and classify thematic funds more quickly and accurately.

In 2019, our Manager Research team identified 1,400 global thematic funds and tagged them across 30-plus themes. This list was created manually by sifting through thousands of research documents over a three-year period. While this process helped us to construct our thematic fund taxonomy, it was also very time-consuming.

To scale our thematic-tagging process, Morningstar created the TRON, Theme Retrieval on Natural Language, model. To our knowledge, this is the first model that applies NLP techniques for thematic fund classification purposes.

---

3 Henry, P., and Krishna, D. "Making the Investment Decision Process More Naturally Intelligent." March 2021. https://www2.deloitte.com/us/en/insights/industry/financial-services/natural-language-processing-investment-management.html

4 Bartram, Söhnke M., Branke, Jürgen, and Motahari, Mehrshad. "Artificial Intelligence in Asset Management." (Sept. 14, 2020). CFA Institute Research Foundation Literature Reviews, August 2020, ISBN 978-1-952927-02-7. SSRN: https://ssrn.com/abstract=3692805 or http://dx.doi.org/10.2139/ssrn.3692805

Many AI models are black boxes. To help explain model outcomes, we also incorporated Explainable AI, an emerging field of AI that aims to improve model transparency, into our framework. We also superimposed human reviews on top of our machine-explained outcomes, which further instills our confidence in the process.

This paper is organized as follows: First, we describe our three-tiered taxonomy and how we built the datasets; second, we discuss how the TRON model efficiently identifies thematic funds; third, we evaluate the model's performance and accuracy; fourth, we demonstrate how we augment model performance by using a Human-in-the-Loop review process; and, finally, we discuss our conclusions and explore future research ideas.

Data: Building the Datasets

In 2019, Morningstar's Manager Research team set out to map the thematic fund universe for the first time. To build a coherent and meaningful definition of thematic investing, analysts combed over more than 50,000 open and closed equity funds[5] in the Morningstar Direct database. Once a yes/no thematic fund tag was assigned to each fund, funds that tracked similar themes were logically grouped together. This was achieved by manually trawling through irregular fund data, including data from fund prospectuses, index methodologies, and other publicly available data sources. This process resulted in a Three-Tiered Taxonomy, seen in Exhibit 1, in which all thematic funds are assigned a Broad Theme, Theme, and Sub-Theme.

5 Our study only includes thematic equity funds. Thematic bond funds were out of scope.

Exhibit 1 Thematic Fund Taxonomy

Technology

Artificial Intelligence + Big Data
Artificial Intelligence
Deep Learning
Quant Computing
Machine Learning
Intelligent Machines
Big Data

Battery Technology
Battery Value Chain
Secondary Battery Industry

Cloud Computing
Cloud Computing

Cyber Security
Cyber Security

Digital Economy
Connectivity
Convergence Technology
Data Economy
Digital Economy
Digitalisation
Disruptive Commerce
E-commerce
Information Revolution
Internet
Internet of Things
New Retail
Sharing Economy
Social Media
Web x.0

Digital Media
Advertising + Marketing Tech
Multimedia

Electronics
Innovative Electronics
Integrated circuits

Fintech
Blockchain
Digital Payments
Disruptive Fintech
Financial Innovation
Fintech
Pay Infrastructure

Future Mobility
Automated Driving
Automobile Innovation
Cleaner Transport
Electric Vehicle
Green Car
Hybrid Car
Internet fo Vehicle
New Car Industry
New Energy Vehicles
Next Gen Auto
Smart Mobiity
Smart Transport

Life Sciences
Alternative Medical Technology
Biotech Clinical Trials
Biotech Innovation
Biothreat
Brand Name Drugs
Cancer
Cardio Devices
Digital Health
Generic Pharma
Genomics
Health Innovation
Ophthalmology
Medical Breakthroughs
Metabolic-Endocrine
mRNA
Neuroscience
Patient Care Services
Treatments Test Advancements

Nano Tech + Smart Materials
Nano Tech
Smart Materials

Next Gen Communications
5G
Advanced Communications
Mobile Internet
Next Generation Communication
Smart Cities
Smart Grid
Telecoms Innovation
Wireless

Robotics + Automation
3D Printing
Automation
Drones
Manufacturing Revolt
Mechatronics
Robotics
Smart Industrial Technology

Space
Space
Space Exploration
Space Innovation

Virtual Reality and Gaming
ESports
Gaming
Online Betting
Virtual Reality

Multiple Technology Themes
4th Industrial Revolution
Disruptive Technology
Early Stage Innovation
Emerging Industry
Exponential Technology
FAANGS
Frontier Tech
Future Economy
IT Revolution
Metaverse
Moonshot Tech
New Economic Engine
Next Generation Technology

Physical World

Energy Transition
Alternative Energy
Carbon Transition
Clean Energy
Clean Tech
Climate Solutions
Decarb Enablers
Green Energy
Hydrogen Economy
New Energy
Nuclear
Renewable Energy
Renewable Infrastructure
Smart Energy
Solar
Wind

Food
AgriTech
Fishing
Food
Food Tech
Nutrition

Logistics + Transportation
E-Commerce Logistics
Jets
Logistics
Ship Building and Transportation
Shipping
Transport

Resource Management
Air
AquaBusiness
Circular Economy
Efficient Resources
Future Resources
New Resources
Rare Resources
Timber
Upstream Natural Resources
Uranium
Waste
Water

Broad Physical World
Multiple Physical World Themes

Social

Cannabis
Cannabis

Consumer
Alcohol
Baby Boomers
Changing Lifestyles
Classic Cars
Consumer Joy
Cosmetics
EM Consumption
Kids
Luxury
Millennials
Organics
Pets
Sports
Subscription Economy

Demographics
Aging Population
Demographics
Future of Humans
Urbanisation

Political
Energy Independence
Infrastructure Spending
Korea New Deal
Korea Unification
New Silk Road
Policy driven
Regional Development
State-Owned Enterprise
Structural Reform
Trade War

Post-Corona
Airlines, Hotels, Cruise Lines
Post Corona
Work From Home
Security
General Security
National Defense
Safety

Security
General Security
National Defense
Safety

Wellness
Alternative Health
Culture
Education
Health + Weight Loss
Healthy Lifestyle
Obesity
Self Development/Fulfillment
Wellbeing

Other
Gig Economy
Holding Companies
Vice

Broad Social
Multiple Social Themes

Broad Thematic

Broad Thematic
Future Life
Future Trends
Global Themes
MegaTrends
New Millenium
Secular Trends
Smart Future
Smart Industries
Transformational Changes

Source: Morningstar Global Thematic Funds Landscape 2022.

While this classification approach allowed us to examine key trends in the thematic fund market, it had a few drawbacks. First, maintaining our thematic fund classification system on an ongoing basis was time-consuming and was limited by analysts' capacities. And second, the human decision-making process could be biased. To mitigate these shortcomings, Morningstar developed an NLP model to improve the efficiency and accuracy of the thematic-tagging process.

Given that the Manager Research team classified funds based on their intentionality, our NLP model uses the same three unstructured data points that analysts examined when they manually classified funds, namely Fund Legal Name, Investment Strategy (English), and KIID Objective/Investment Policy. The definitions of these data points can be found below.

Fund Legal Name: A fund's legal name. This will be referred as FN in the exhibit below.

Investment Strategy (English): The first sentence will always be the fund's investment objective. From there, the rest of the description will be a summary of that fund's principal investment strategies as written in the prospectus—this should first include what a fund "normally" or "primarily" invests in, followed by what the fund "may" invest in. Additionally, it includes information about what the fund does not invest in, if applicable. Finally, if the fund is nondiversified, it will include a nondiversification statement. This will be referred as IS in the exhibit below.

KIID Objective/Investment Policy: The objective of the fund as stated in the KIID document. This will be referred as KO in the exhibit below.

Our universe consists of around 65,000 open-end funds and ETFs of equity asset class. Additionally, the manually curated list of 1,400 thematic funds served as our Ground Truth. We used this list to evaluate the efficacy and accuracy of our machine-learning model. In the next section, we describe our approach to model selection.

Model Selection: Supervised and Unsupervised Learning

Most classification problems can be modeled using either a supervised or an unsupervised learning approach. Supervised learning models aim to map an input to an output using a labeled training data set, which comprises input and output pairs. A supervised learning algorithm analyzes the training data and learns/determines an empirical function, which can then be used to classify unlabeled data. An optimal model will correctly determine the class labels from unseen instances. The most widely recognized model that follows this approach is linear regression.

By comparison, the unsupervised approach can be used when there are no output variables—in our case this is thematic labels. In this instance, we cannot supervise the model's learning. Instead, we seek to understand the relationship between the input variables. One way we can examine this relationship is by using clustering. If observations share similar characteristics, we may be able to classify them into distinct categories.[6]

Unsupervised learning models can be used to uncover patterns in the data without knowing the underlying themes. The hope is that through sufficient training the model can identify an internal representation of underlying data and then generate themes from it. The unseen representation can then be generalized to detect similarities and differences between themes.

Given that Morningstar's Manager Research analysts manually classified global thematic funds, we already had a labeled dataset. As such, we began our model selection process by considering supervised

---

6 Witten, D., Hastie, T., Tibshirani, R., and Gareth, J. "An Introduction to Statistical Learning: With Applications in R (Springer Texts in Statistics)." New York, London: Springer, 2013.

learning models, such as logistic regression and random forest. While the model initially showed promising outcomes, there were some inherent challenges with the learning process.
First, our fund classification system was imbalanced across all themes. Our Ground Truth dataset of 1,400 manually classified funds was unevenly split across 30 themes (classes). For example, themes like Electronics and Post-Corona were underrepresented in the dataset. All told, these models did not perform because there was not enough training data per class.

Second, the size of the training set was small at around 1,400 funds. By comparison, Morningstar's fund universe spans over 50,000 funds. As such, the dataset we used to train the model was not representative of the entire fund population that we aimed to tag eventually. Given these limitations, we decided to try alternative approaches.

Morningstar's TRON Model Explained

To replicate our analyst-driven thematic-tagging process, we built an NLP model, which uses relevant text information to classify funds into their respective themes. Once themes are assigned to funds, we use business logic rules to classify funds into broad themes. Our model currently does not classify funds into sub-themes.

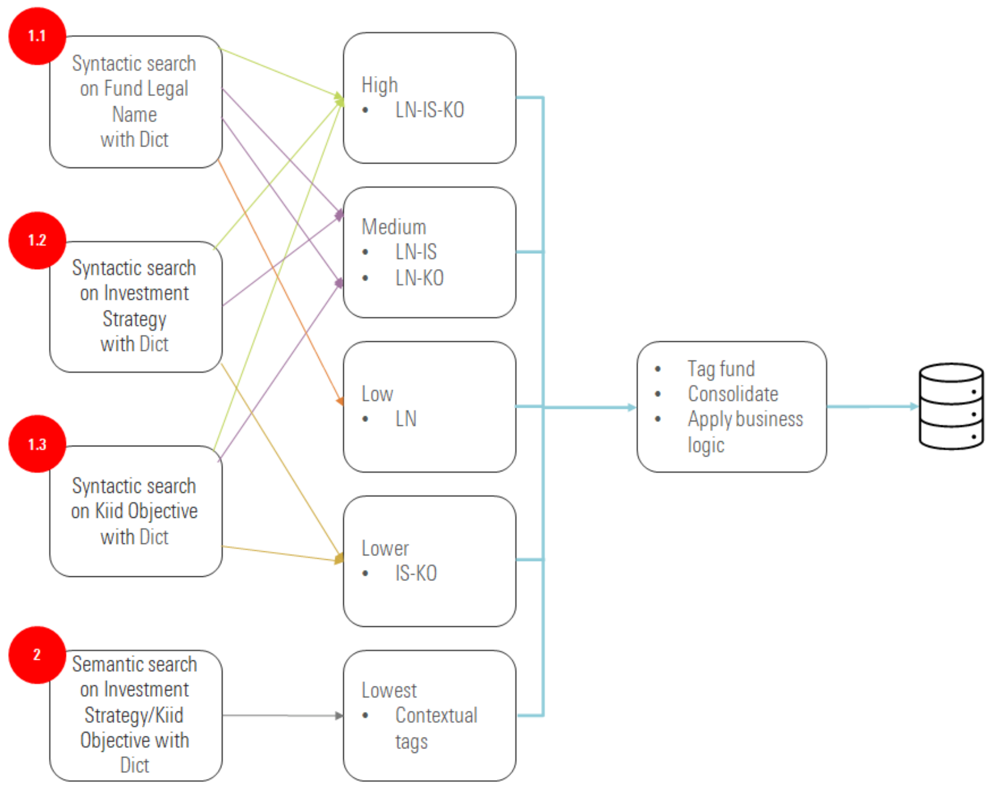
Our proprietary TRON model classifies funds based on their intentionality, rather than holdings. We followed this approach to replicate our analyst-driven thematic fund classification process. One drawback of this approach is that it assumes fund management companies truthfully disclose information on their funds' intentions.

Our NLP algorithm searches for thematic text information syntactically and semantically. Our Syntactic Search uses a text-search algorithm that detects the occurrence of keywords that are associated with a particular theme in a fund's name, investment strategy text, and KIID objective. This approach looks for literal matches of the keywords or variants of them, without understanding the meaning. In our dictionary, the key is the theme, and the values are the thematic keywords that were identified by our analysts. Given that the thematic industry is evolving with baffling rapidity, these key-value pairs are maintained on an ongoing basis.

Meanwhile, the Semantic Search quantifies and categorizes semantic similarities between linguistic items. This approach determines how close two texts are in meaning. It captures deeper contextual meanings. For example, while "How old are you?" and "What is your age?" are grammatically different, the meaning of the two questions is semantically similar. In our case, it can recognize that, "Marijuana" is semantically similar to "Cannabis". Finally, we assign a high, medium, or low confidence score to our predicted tags.

Exhibit 2 provides a visual overview of the process. Here, Steps 1 and 2 are sequential, whereas sub-steps 1.1, 1.2, and 1.3 are parallel. For example, if the NLP model detects relevant thematic information in a fund's documentation for a theme simultaneously in its name, investment strategy, and KIID Objective, then our model will assign a high confidence score to that theme. In contrast, if we find relevant information only in a fund's name, investment strategy, or KIID Objective, then the model will have a medium or low confidence score for that theme.

Exhibit 2 TRON - Theme Retrieval on Natural Language



Source: Morningstar Data. March 2022. Dict is the dictionary that contains the thematic taxonomy.

We assign Broad Themes based on business logic rules. Our model tagged Global X Robotics & Artificial Intelligence ETF to the Themes - Robotics + Automation and Artificial Intelligence + Big Data, which was then assigned a final Broad Theme, such as Multiple Technology Themes due to a business logic (Funds tagged to more than one technology Theme should be assigned Multiple Technology Themes).

In the next section, we evaluate the model's performance metrics.

TRON Model Evaluation

We measured the TRON model's performance by comparing the predicted tags with analyst-assigned thematic labels, which served as our Ground Truth.

Accuracy is calculated as the sum of the number of correct predictions divided by the total number of predictions.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)$$

Where,

TP = True Positives
PN = True Negatives
FP = False Positives
FN = False Negatives

We measured the performance by three metrics, namely the weighted Precision, Recall, and F1-scores. We also weight these metrics by the number of observations/funds per class/themes.

Precision is the number of true positive predictions divided by the sum of true positive and false positive predictions. Precision measures what proportion of positive identifications are correct. A higher Precision value indicates that the false positive rate is low. Our weighted precision score for all themes is 96%, which is on the higher side.

$$Precision = tp/(tp + fp)$$

Recall represents the true positive rate. The number of true positive predictions is divided by the sum of the true positive and false negative predictions. Recall measures the proportion of true positives that were identified correctly. A model that produces no false negatives has a Recall of 1. Our weighted Recall score for all themes is 81%. This indicates that we were able to identify most of the funds that the analyst team tagged using the TRON model.

$$Recall = tp/(tp + fn)$$

Finally, the F1-score combines Precision and Recall into a single metric by taking the harmonic mean of Precision and Recall. An F1 score that is closer to 1 means that both Precision and Recall are high. Meanwhile, an F1 score that is closer to 0 indicates that both Precision and Recall are low. An F1 score that is close to 0.5 means that one metric is high while the other is low. Our model's F1-score stands at 0.87, which is relatively close to 1. This also means that our model correctly labeled 87% of the funds that were tagged by analysts.

$$F1\ Score = 2 * (Precision * Recall)/(Precision + Recall)$$

Exhibit 3 below shows the TRON model's performance metrics for individual themes. A weighted Precision, Recall, and F1 score are also calculated for all themes.

Exhibit 3 Model Performance

-	precision	recall	f1-score	support
Artificial Intelligence + Big Data	1.00	0.87	0.93	52
Battery Technology	1.00	0.78	0.88	9
Broad Thematic	0.96	0.70	0.81	152
Cannabis	1.00	0.88	0.94	26
Cloud Computing	0.94	1.00	0.97	16
Consumer	0.97	0.86	0.91	184
Cyber Security	1.00	0.74	0.85	19
Demographics	0.83	0.77	0.80	39
Digital Economy	0.97	0.88	0.92	142
Electronics	0.67	1.00	0.80	2
Energy Transition	0.99	0.75	0.85	172
Fintech	0.98	0.88	0.93	52
Food	0.86	0.86	0.86	14
Future Mobility	0.97	0.63	0.76	59
Life Sciences	0.97	0.73	0.83	92
Logistics + Transportation	0.91	0.95	0.93	22
Nano Technology + Smart Materials	0.86	1.00	0.92	6
Next Gen Communications	0.97	0.79	0.87	42
Political	0.95	0.78	0.86	51
Post-Corona	1.00	1.00	1.00	5
Resource Management	0.99	0.93	0.96	106
Robotics + Automation	0.91	0.82	0.86	60
Security	0.73	0.64	0.68	25
Space	1.00	1.00	1.00	8
Virtual Reality and Gaming	1.00	0.85	0.92	26
Wellness	0.81	0.89	0.85	28
accuracy			0.81	1412
weighted avg	0.96	0.81	0.87	1412

Source: Morningstar Research.

Human-in-the Loop Review Process

TRON has a Human-in-the Loop review process. The model's predicted thematic fund tags are reviewed by our global team of subject-matter experts. We have decided to take this approach because thematic fund classification is not always a clear-cut exercise.

The line between sectors and themes can be blurry, especially as sectors' definitions have drifted over time. Perhaps the most challenging distinction to make is between tech-sector funds and those that track one or more tech-related themes. Differentiating between sustainable and thematic funds can also prove challenging. For example, thematic funds in the Energy Transition Theme tend to also have environmental, social, and governance attributes.

As thematic fund classification can be widely debated, we have decided to impose two levels of analyst reviews. The first review occurs at the regional level. More specifically, Morningstar's analysts in Europe, Asia, and the Americas review thematic tags assigned to funds domiciled in their respective regions. To ensure business continuity, we have a primary and a backup analyst in each region.

The second review is conducted by one of our thematic fund methodology owners. These subject-matter experts have a more in-depth understanding of our thematic fund taxonomy. The final reviewer can either accept the regional reviewer's tag and update our thematic fund universe or reject the proposed tag and reclassify the fund.

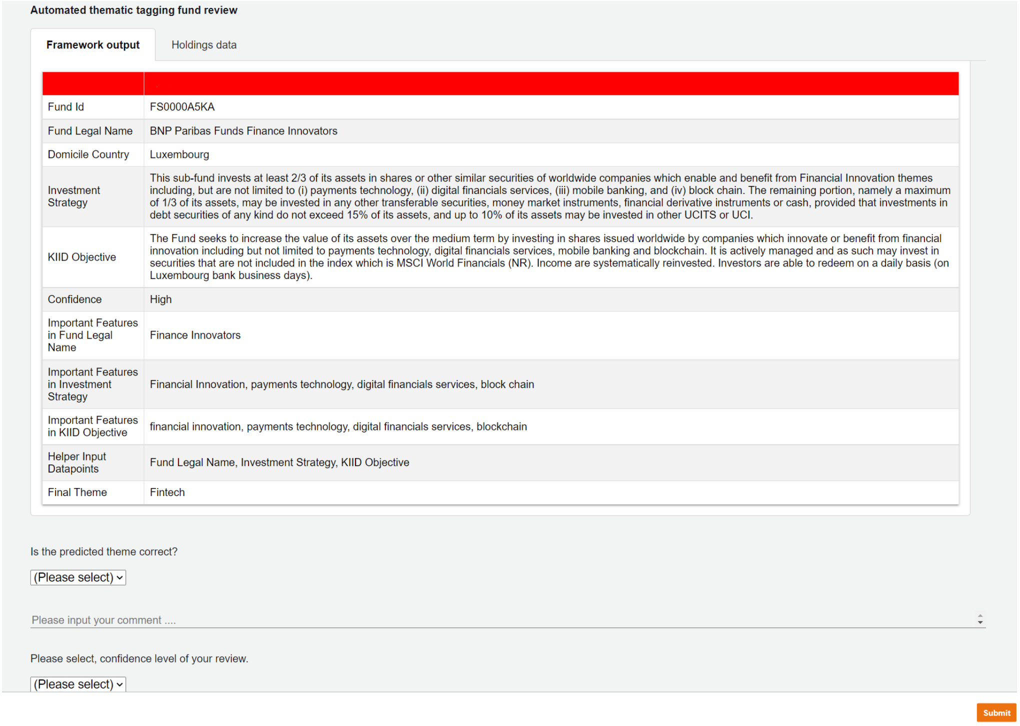
Our infrastructure allows analysts to add comments to thematic fund tags. As such, each decision has an audit trail. The two review levels also ensure we have a robust governance process in place, allowing us to reach consensus on how to tag new funds, evaluate new themes, and critique model outcomes. Analysts are also able to evaluate TRON's predictions by looking under the hood of the model. In the next section, we discuss how we address model explainability.

TRON Model Explainability

As we discussed in the introduction, we cannot solely rely on black box outcomes generated by machine-learning models, as they may suffer from inherent data biases. To address this issue, we provide analysts with information on why the model assigned a specific theme to a fund as part of our Human-in-the-Loop review.

The exhibit below shows how we address model explainability. In this example, the TRON model assigned the Fintech Theme to the BNP Paribas Funds Finance Innovators fund with a high confidence level.

Exhibit 4 Human-in-the-Loop User Interface


Automated thematic tagging fund review

Framework output Holdings data

Fund Id	FS0000A5KA
Fund Legal Name	BNP Paribas Funds Finance Innovators
Domicile Country	Luxembourg
Investment Strategy	This sub-fund invests at least 2/3 of its assets in shares or other similar securities of worldwide companies which enable and benefit from Financial Innovation themes including, but are not limited to (i) payments technology, (ii) digital financials services, (iii) mobile banking, and (iv) block chain. The remaining portion, namely a maximum of 1/3 of its assets, may be invested in any other transferable securities, money market instruments, financial derivative instruments or cash, provided that investments in debt securities of any kind do not exceed 15% of its assets, and up to 10% of its assets may be invested in other UCITS or UCI.
KIID Objective	The Fund seeks to increase the value of its assets over the medium term by investing in shares issued worldwide by companies which innovate or benefit from financial innovation including but not limited to payments technology, digital financials services, mobile banking and blockchain. It is actively managed and as such may invest in securities that are not included in the index which is MSCI World Financials (NR). Income are systematically reinvested. Investors are able to redeem on a daily basis (on Luxembourg bank business days).
Confidence	High
Important Features in Fund Legal Name	Finance Innovators
Important Features in Investment Strategy	Financial Innovation, payments technology, digital financials services, block chain
Important Features in KIID Objective	financial innovation, payments technology, digital financials services, blockchain
Helper Input Datapoints	Fund Legal Name, Investment Strategy, KIID Objective
Final Theme	Fintech

Is the predicted theme correct?

(Please select)

Please input your comment ....

Please select, confidence level of your review.

(Please select)

Submit

Our interface also shows relevant text information like Fund Name, Investment Strategy, and KIID Objective. We also provide data on which key words contributed the most to the final prediction. Finally, our interface reveals which sections of the documents were identified as relevant by our Syntactic and Semantic searches. This infrastructure enables us to better explain why themes were assigned to funds. It also allows us to maintain an audit trail of our model's outcomes.

The Human-in-the-Loop cycle occurs monthly, ensuring that our Ground Truth gets updated to reflect the latest thematic fund universe. Also, any funds that are detected as non-thematic are flagged and prevented from popping up again in the analysts' review process. The model's outcomes are periodically reviewed and amended if we observe diminishing accuracy. Finally, the entire process is tightly guarded by a set of quantitative analysts, ensuring watertight control over the quality of our generated tags.

Concluding Thoughts

In this paper, we described our novel approach to tagging thematic funds. Our approach combines machine intelligence with Explainable AI and multiple Human-in-the-Loop reviews. As such, the model is transparent, understandable, and comprehensive.

This approach has enhanced our process in several ways. First, the model has expanded our coverage of global thematic funds. Under the analyst-driven approach, we were able to tag 1,400 global thematic funds. By comparison, we tagged 2,500-plus funds using the new NLP and Human-in-the-Loop process, which, in turn, became the new Ground Truth.

Second, the revamped model reduced analysts' manual intervention, which then increased our business efficiency and capacity. And third, the new approach has made our process more flexible—it allows us to quickly scan the entire fund universe and generate new thematic fund tags when a new theme is detected.

Our engine currently supports documents written in English. We are looking to expand our capabilities to support more languages. We also wish to uncover new themes. Currently, our algorithm tags funds based on a text-search algorithm. We are also looking to incorporate holdings-based data into the classifier. Finally, we also plan to extend the framework to other potential use cases like ESG and Strategic Beta.

References

"How Much Data Do We Create Every Day? The Mind-Blowing Stat Everyone Should Read," by Bernard Marr, Forbes.com, March 21, 2018.

"Making the Investment Decision Process More Naturally Intelligent," by Patrick Henry and Dilip Krishna, Deloitte.com, March 2, 2021.

"Machine Learning in UK Financial Services," Bank of England, October 2019.

"AI Pioneers in Investment Management: An examination of the trends and use cases of AI and big data technologies in investments," CFA Institute, 2019.

About Morningstar Quantitative Research

Morningstar Quantitative Research is dedicated to developing innovative statistical models and data points, including the Morningstar Quantitative Rating, the Quantitative Equity Ratings, and the Global Risk Model.

About Morningstar Manager Research

Morningstar Manager Research provides independent, fundamental analysis on managed investment strategies. Analyst views are expressed in the form of Morningstar Analyst Ratings, which are derived through research of five key pillars—Process, Performance, Parent, People, and Price. A global research team issues detailed Analyst Reports on strategies that span vehicle, asset class, and geography. Analyst Ratings are subjective in nature and should not be used as the sole basis for investment decisions. An Analyst Rating is an opinion, not a statement of fact, and is not intended to be nor is a guarantee of future performance.

For More Information

+1 312 244-7541
lee.davidson@morningstar.com

MORNINGSTAR®

22 West Washington Street
Chicago, IL 60602 USA