

Sentiment analysis for financial services use cases



Assessing sentiment for equity trading, credit markets, and customer experience.

NetApp and SFL Scientific have combined their expertise to help customers address important AI use cases quickly. In the financial services industry, the ability to accurately assess the sentiment of spoken or written communications has significant benefits for diverse use cases, including equity trading, assessing credit markets, and understanding the customer's experience. This paper describes the rapid implementation of a state-of-the-art deep learning model to detect sentiment in diverse spoken and written communications. The general methodology described is applicable to a broad range of NLP and other AI challenges.

Competition in the financial services industry has increased significantly in recent years, with tech giants and fintech startups putting significant pressure on industry incumbents. Success today often hinges on the ability to take full advantage of all available data. There is no shortage of voice and text data, but it can be a challenge to reliably extract clear signals from this data that might improve the customer's experience or underpin a trading or investment strategy. The overwhelming volume of data makes it difficult to identify clear trends that can inform decision making or enable new algorithms. As a result, this data may be underused as a source of insight.

Sentiment analysis offers opportunities to uncover meaningful insights embedded in quarterly earnings calls, research reports, company filings, and other sources. A [recent study in Nature](#) found that “there exists a weak but statistically significant association between strong media sentiment and abnormal market return as well as volatility.” Using natural language processing (NLP), sentiment analysis can be used to rapidly analyze large quantities of voice and text data, delivering insights before the currency and value of the data decline.

NetApp and SFL Scientific have developed an easy-to-implement AI pipeline that captures sentiment from text or audio, including real-time conversations. Our solution can be quickly deployed on premises or in the cloud and trained and tailored to your specific use case to extract valuable insights and improve your company's competitive position.

An AI pipeline for sentiment analysis in financial services

AI modeling techniques can be used to perform nuanced analysis of a wide range of voice and text data, such as quarterly earnings reports, social media, news media, and many others. We have created a simple AI pipeline that identifies the sentiment of spoken language or text on a sentence-by-sentence basis. This granularity makes it possible to score and track sentiment, including identifying inflection points and important junctures in text or conversation.

Tracking sentiment can provide an objective assessment of sentiment in published text and during interactions between companies, institutional investors, and analysts.

By combining NetApp's hardware and data management tools with SFL Scientific's comprehensive data science and data engineering skills, we developed an end-to-end AI pipeline that is applicable to financial services use cases. The pipeline uses a modular architecture, can be deployed quickly, and can be continuously refined to improve accuracy and satisfy the needs of your business.

In operation, voice and text data is processed through a pipeline with pretrained speech recognition and sentiment modules to determine sentiment at the sentence level. Audio data is ingested and converted to text, then each sentence is assigned a sentiment value of positive, negative, or neutral. Sentiment is scored per sentence and tracked over time, with results sent to a dashboard or stored for subsequent analysis. Results are available in near real time for use cases that require it. For example, if sentiment analysis shows that an earnings call turns negative during Q&A, that information could determine the next day's trading strategy for the company's stock.

Pipeline components

The pipeline is built using NVIDIA software running on [NetApp® ONTAP® AI](#) hardware to facilitate training and inferencing. NetApp ONTAP AI combines the capabilities of [NVIDIA DGX systems](#) and NetApp AFF to remove performance bottlenecks and accelerate AI workloads, and the [NetApp DataOps Toolkit](#) streamlines data storage and handling. Our pipeline can ingest multiple streams of data in real time and pass it on for transcription (in the case of voice data) and sentiment analysis.

The pipeline can optionally be architected to incorporate on-premises or cloud-based object storage to accommodate the needs of distributed or cloud-based use cases and for long-term archiving, as shown in Figure 1.

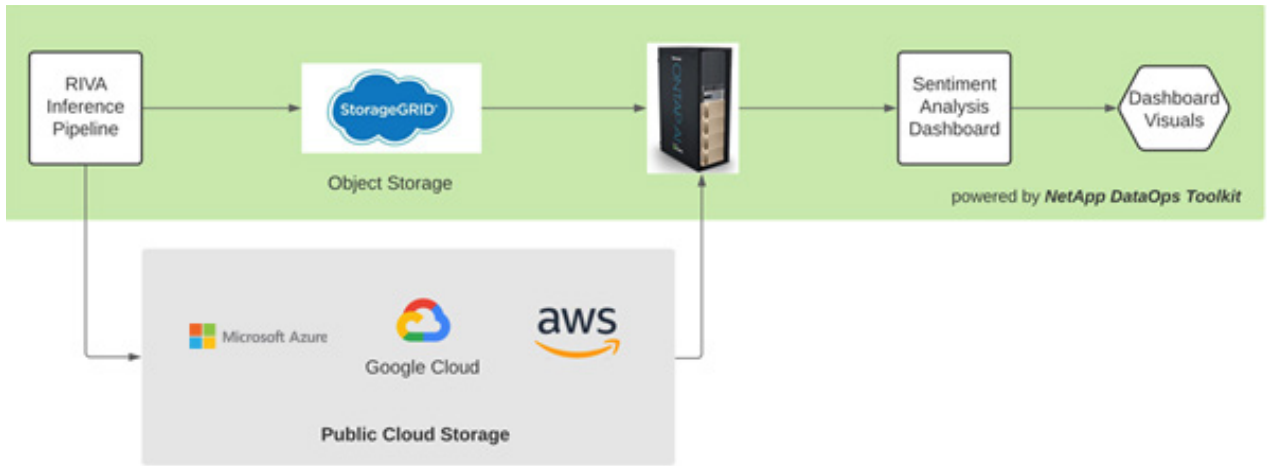


Figure 1) Object storage can be incorporated into the pipeline to accommodate distributed or cloud-based use cases and to provide a long-term archive of sentiment data for trend analysis.

Audio transcription

Speech-to-text conversion (when needed) is accomplished by using automated speech recognition (ASR). Each audio stream is preprocessed, including conversion to WAV format and separation into sentences, and passed to [NVIDIA Riva](#), a GPU-accelerated software development kit (SDK) for NLP. The Riva server returns a text representation of each audio sentence and passes it on for sentiment analysis.

Sentiment analysis

NVIDIA Riva is also capable of classifying text in various ways. This project uses a pretrained Riva module to detect sentiment, eliminating the up-front training effort necessary to train an AI model from scratch. The module classifies the sentiment of each sentence as neutral, positive, or negative, providing near-real-time insight. The end-to-end pipeline is illustrated in Figure 2.

Proof of concept

To assess the functionality and performance of this pipeline, we needed a real-world audio dataset that could be used to test both our audio transcription and sentiment modules. We settled on a publicly available dataset containing thousands of earnings call audio recordings and transcripts for S&P 500 companies. This dataset is representative of the many types of communication prevalent in the financial services industry and covers a wide range of industry sectors.

In testing, our pipeline is capable of processing hundreds of such calls in parallel. Because this dataset includes transcriptions, it is ideal for training and fine tuning the speech recognition and text classification models. Using this dataset, updated sentiment analysis appears on the output display less than 1 second after a sentence is ingested, demonstrating its ability to support near-real-time use cases.

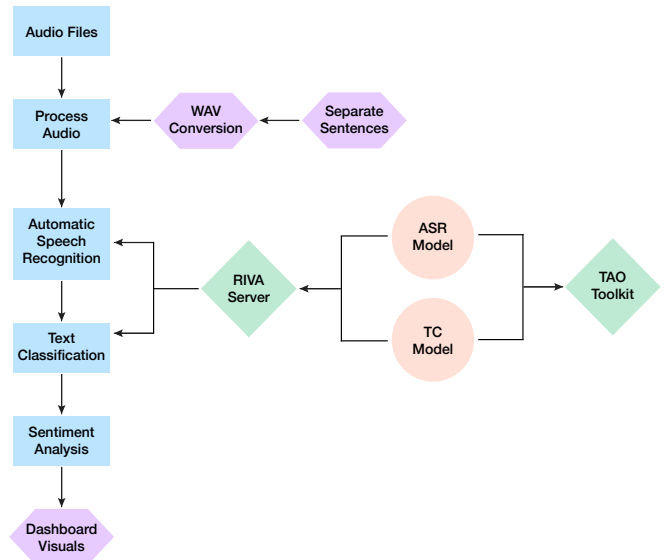


Figure 2) Pipeline to determine sentiment from call center audio.

Transfer learning and fine tuning

The [NVIDIA NGC catalog](#) includes many pretrained models for use with the Riva SDK. These models provide a starting point for pipeline optimization that can be quickly fine tuned to meet specific financial services needs by using transfer learning.

[Transfer learning](#) is the process of fine tuning a previously trained neural network for a similar or new use case. NetApp and SFL Scientific start with pretrained deep learning models and can quickly fine tune them for your needs by training with a set of data that is specific to your desired use case.

Typically, our data scientists experiment with a range of model configurations and maximize model performance by using multiple model tuning experiments. They make adjustments during each iteration and select the model that performs the best.

Training and data management tools

The NetApp DataOps Toolkit and the [NVIDIA Train, Adapt and Optimize \(TAO\) Toolkit](#) are used with this solution to streamline transfer learning and fine tuning. Using these tools to operate and maintain unified data pipelines reduces complexity, enables teams to efficiently manage data, and ensures traceability.

[The NetApp AI Control Plane](#) pairs machine learning operations (MLOps) with NetApp technology to simplify the management of AI data and to facilitate experimentation. The NetApp DataOps Toolkit makes it easier to manage the large volumes of voice and text data that financial services companies may want to analyze.

Used separately or together, these tools significantly speed up AI projects, enabling you to set up and clone the volumes needed for training, evaluate results, and iterate quickly, ensuring reproducibility, reliability, and compliance.

Retraining

You can periodically retrain and fine tune your models by annotating text from existing transcripts to better match past findings. NetApp and SFL Scientific can help you understand how to manage and annotate data to optimize your models for the best possible results. This annotated data can be fed through the training pipeline to keep your AI models and pipeline tuned to your needs.

Sentiment analysis use cases for financial services

Our sentiment analysis model can be tailored to address a broad array of FSI use cases, including the following.

- **Tracking equity markets.** The pipeline can be used to track sentiment from news media, social media, press releases, and so on to identify companies that are regarded positively—or negatively—as a guide to financial trading. Additional correlations can be performed with other market signals as part of advanced trading algorithms. For instance, you may want to correlate public sentiment with stock performance and other data to predict the performance of a particular stock, a set of stocks, or an industry sector.
- **Monitoring credit markets.** A related use case is to apply sentiment analysis to credit markets such as corporate bonds, government bonds, collateralized debt obligations (CDOs), and credit default swaps (CDSs). Sentiment can be gathered from a variety of sources to monitor a particular credit market or individual securities.
- **Customer sentiment.** Social media and the internet capture data about your products, services, and brand that may be the truest indicator of how customers feel about you—and your competitors. Our pipeline can be applied to analyze the sentiment of social data to gauge the success of products and services and the relative strength of your company versus competitors. Are people positive or negative about your company, brand, product, or service? What are they saying about your competitors? Is the trend in sentiment positive or negative? What was the impact of the latest product rollout or the latest financial report? The resulting data can help you guide future projects, better address customers' desires, and improve your brand.

Additional AI opportunities

In addition to the use cases just described, sentiment analysis tools are applicable to a wide variety of enterprise use cases, including call centers. Looking beyond sentiment, the same approach—combining transfer learning, rapid experimentation, iterative fine tuning, intelligent data management, and production deployment with regular retraining—can be applied to NLP and other use cases to address a variety of financial services needs.

NetApp and SFL Scientific get your AI project to production more quickly with fewer missteps. If you have the data, we can help you gain the insight your business needs.

About our partnership

The partnership between NetApp and SFL Scientific brings together SFL's proven data science and data engineering expertise and NetApp's industry-leading AI hardware and software.

Get started

To find out whether NetApp and SFL Scientific can help you accelerate your path to AI, visit netapp.com/ai and fill out the form to request a contact.

NetApp

Overcoming the challenges of building an AI architecture for machine learning (ML), deep learning (DL), and other AI workflows requires careful infrastructure planning. [NetApp](#) helps customers build AI pipelines that smooth the flow of data and deliver optimal results.

NetApp offers a complete portfolio of AI hardware and software solutions and services to get your data flowing from edge to core to cloud. netapp.com/ai



SFL Scientific is a data science consulting and professional services company operating on the highest end of technology development. We specialize in data engineering, machine learning, Artificial Intelligence, and big data, continuously consolidating a pool of the very best data science and analytics talent for our clients. www.sflscientific.com



+1 877 263 8277

About NetApp

In a world full of generalists, NetApp is a specialist. We're focused on one thing, helping your business get the most out of your data. NetApp brings the enterprise-grade data services you rely on into the cloud, and the simple flexibility of cloud into the data center. Our industry-leading solutions work across diverse customer environments and the world's biggest public clouds.

As a cloud-led, data-centric software company, only NetApp can help build your unique data fabric, simplify and connect your cloud, and securely deliver the right data, services and applications to the right people—anytime, anywhere. To learn more, visit www.netapp.com



© 2021 NetApp, Inc. All Rights Reserved. NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners. CSS-7217-1221