Developing a Confidence Scoring System for Business Data Validation in AI Systems

By

David Sackler

William Wang

Dev Patel

Esteban Felix Tapia

Supervisor: Abid Ali

A Capstone Project

Submitted to the University of Chicago in partial fulfillment
of the requirements for the degree of

Master of Science in Analytics

Division of Physical Sciences

August 2025

Abstract

In today's business landscape, verifying the legitimacy of companies is often hindered by inconsistent or incomplete data, posing challenges for analysts and businesses in assessing reliability. BrightQuery, a provider of historic and current financial information, aims to leverage its data to generate a confidence score for businesses worldwide. The confidence scoring system would evaluate the reliability of business data, using key factors such as legal entity validation, financial reporting, and online presence. This system integrates BrightQuery's financial, employment, and payroll data to assess company reliability, employing scoring algorithms and visualizations to communicate confidence scores effectively.

Keywords: Business Legitimacy, Data Integrity, Confidence Scoring, AI-driven Risk Assessment, Data Validation, Company Reliability

Executive Summary

In today's data-driven landscape, companies wrestle with the complex task of verifying the legitimacy of other businesses. BrightQuery confronts this challenge head-on by engineering a confidence system that evaluates the reliability of businesses in the United States through critical indicators like legal entity validation, financial reporting and online presence. This system empowers clients with a robust and nuanced view of business legitimacy, sharpening both risk assessment and strategic decision-making capabilities.

Drawing from authoritative sources such as IRS and Department of Labor filings, BrightQuery curates comprehensive financial and employment profiles for over 74 million businesses. The dataset spans business demographics, payroll statistics, and core financial metrics. This project aims to construct a global dynamic confidence scoring system that distills complex data into clear, interpretable confidence scores—both absolute and comparative—delivering valuable insights into a company's trustworthiness.

The scoring methodology utilizes confidence scores, generated by BrightQuery's US model, to label the US businesses within the global dataset. Supervised techniques applied to this subset of the global data learn the relationship between the global features and confidence scores. The trained model is then applied to every business in the global dataset, generating confidence scores with the global data. The system generates transparent, data-driven scores for each business, equipping users with a deeper understanding of organizational credibility and risk.

Table of Contents

List of Figures

List of Tables

Introduction

BrightQuery exclusively provides comprehensive historical and current financial, employment, and payroll data for private and public U.S. companies, sourced from validated, up-to-date government filings. Covering millions of businesses, BrightQuery will systematically score the reliability of businesses based on their financial and employment data. This system helps clients gain deeper insights into business credibility, supporting better risk management and investment decisions.

Opportunity Statement

Currently, BrightQuery's scoring system is focused on US-based companies, but there is growing demand for a global solution. This presents an opportunity to evolve the platform—expanding its capabilities to generate consistent and standardized confidence scores for businesses worldwide. The goal is to build a more objective and context-sensitive scoring system that delivers both absolute and relative scores, helping clients across global markets assess business legitimacy and credibility with greater accuracy.

This opportunity positions BrightQuery to develop a powerful framework that enables clients to validate business legitimacy, detect fraud, pursue secure investments, and engage in partnerships with greater confidence. By analyzing key indicators—such as legal entity validation, financial reporting, and online presence—the system delivers comprehensive scores that reflect a company's credibility and reliability from multiple perspectives. With these insights, clients—estimated by BrightQuery to include around 3,000 companies—can make more accurate, data-informed decisions while significantly minimizing risk. Whether assessing creditworthiness, evaluating lending opportunities, or making underwriting calls, this tool becomes an essential asset in streamlining and strengthening risk assessment processes.

Analysis Goals

The goal of this analysis is to develop a confidence scoring system that accurately reflects a company’s reliability and credibility based on its data. This system will integrate key factors such as financial reporting, legal entity validation, and online presence into a comprehensive score that assesses business legitimacy.

To achieve this, we will utilize the provided data as our foundation and actively analyze a sample of US businesses within the global dataset labeled by BrightQuery’s existing model. Success will be defined by our ability to achieve a low RMSE of below 15 relative to the established labeled scores. Additionally, the system will feature both relative and absolute scores—relative scores for companies of similar origin and industry, and absolute scores for all companies, allowing for more precise comparisons and assessments.

A key component of this analysis is the creation of a user-friendly dashboard or visualization tool. This will allow users to view and compare confidence scores across different companies and industries, rank businesses based on various metrics, and understand how the individual data points contribute to the overall score. By visualizing these factors, users will gain deeper insights into a company’s credibility.

Scope

The scope of this project is focused on developing a confidence scoring system that evaluates a company's reliability and credibility using validated data from financial, employment, and payroll records sourced from government filings. The scoring model will specifically assess key factors such as legal entity validation, company type, financial reporting, and online presence. However, the project will not consider external factors like customer sentiment, market reputation, or global business regulations, which could impact a company’s overall legitimacy. Additionally, the data used in this system will be confined to structured data stored in SQL tables, excluding unstructured data sources such as customer reviews, social media sentiment, or other external datasets.

The project will concentrate on creating a scoring algorithm and an interactive dashboard for visualizing the confidence scores. It will not address scoring for companies based on data outside of what is provided in the specified structured dataset. Additionally, the extraction and integration of data from other external sources for further scoring will not be part of this project's scope. These limitations will help focus the development on delivering a reliable, data-driven scoring system within the defined parameters.

Background

In developing a robust confidence scoring system to assess business legitimacy, we draw upon a range of proven methodologies from credit risk modeling, business evaluation, and data-driven financial analysis. Across the literature, one consistent theme emerges: a company's financial health—particularly its profitability, debt levels, and ability to manage cash—is crucial in determining its stability and trustworthiness over time. Key financial indicators that we have seen across the literature and risk scoring models include the following:

1. EBITDA (Earnings Before Interest, Taxes, Depreciation, and Amortization): a measure of a company's core profitability before accounting for financial or tax-related costs.
2. Leverage ratios: metrics that show how much debt a company is using to finance its operations compared to its equity or assets.
3. Cash flow ratios: financial metrics that use information from a company's statement of cash flows—particularly from operating activities—to assess its financial health. **Cash flow** refers to the actual cash a company receives and spends over a specific period. It shows how much cash the company collects from customers for its products or services and how much cash it pays out for everyday expenses like salaries, suppliers, and bills.
4. Debt coverage: a company's ability to pay back its debts using its income.
5. Growth metrics: indicators of how quickly a company is expanding, such as revenue or asset growth over time.

Literature Review

Business validation is a field that involves estimating a subjective score, a score that represents the validity of a business. The term "business value" in this context is not well-defined, as individuals may prioritize different factors when evaluating legitimacy. For instance, some may consider company size more important than cash flow, while others may view the reverse as true. This subjectivity highlights the importance of designing a scoring system that mitigates bias, particularly size bias, and presents results through a user-friendly dashboard. Our approach leverages well-documented research in this area to guide the development of our confidence scoring system.

The earliest forms of business validation started in medieval Europe with the merchant guilds. These organizations played a crucial role in establishing trust in commercial transactions and protecting merchant rights. Guilds acted as a mechanism called “intertransactional linkages” (Greif, Avner), which helped “change the set of self-enforcing beliefs in the ruler-trader transaction” (Greif, Avner). They emerged as a response to commitment problems between different traders and merchants. The collective enforcement provided by the merchant guilds offered security for merchants and their goods, effectively introducing the idea of business legitimacy verification.

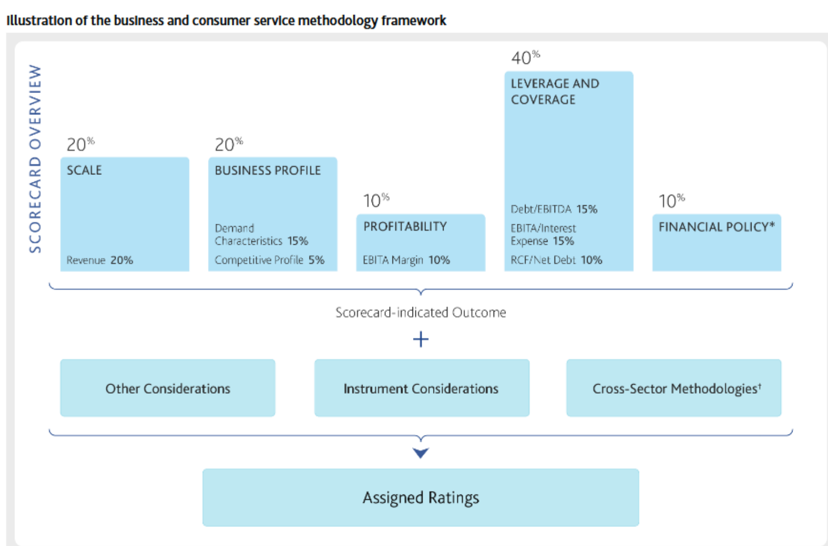
Official licensing appeared between the 12th and 16th centuries, when guilds received privileges to exclusively perform their trade. These privileges came in the form of charters and are recognized as the earliest recorded form of licensing. These charters were primarily used to monopolize a trade and generate revenue (Davron).

BrightQuery’s competitors, such as Moody’s and Middesk, have pioneered approaches to generating a business validity score. These companies have established methodologies that assess credibility based on a variety of features and risk indicators.

A common method for quantifying a business’s legitimacy is through a linear regression model, which assigns weights to features to assess their relative impact. Dwyer, Douglas W., et al. (2004) introduced the Moody’s KMV RiskCalc™ v3.1 Model, which incorporates features such as the cash flow ratio, debt coverage, and liquidity to quantify credit risk. One of the strengths of this model is its effective mitigation of size bias through the use of cash flow ratios, which standardize comparisons across firms of varying scales. Additionally, the model integrates industry-specific indicators to embed market insights at the sector level, avoiding misleading direct comparisons between private firms and publicly implied market values. These aspects, bias mitigation, industry-level insights, and effective feature selection, directly inform our research and modeling strategy.

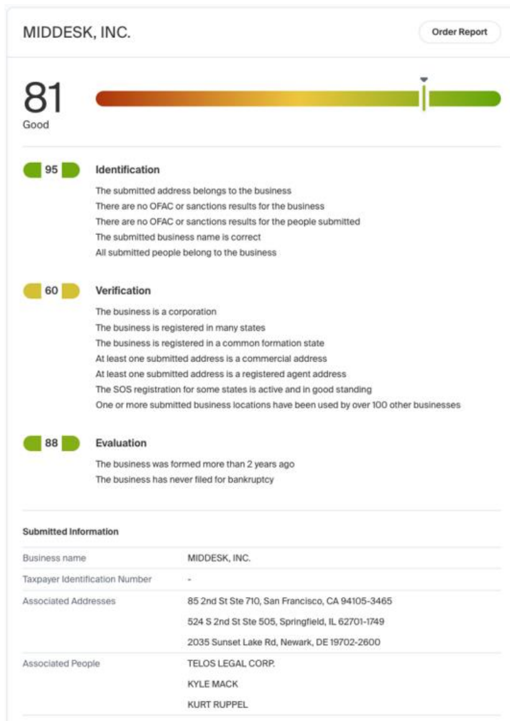
Moody's has also developed a scorecard system (Moody's Investors Service) as shown in figure 1 that expands on this methodology, outlining both the weights applied to business characteristics and the rationale behind those choices. Their framework demonstrates a structured, transparent way to assess business creditworthiness.

Figure 1. *Illustration of Moody's Scorecard System*



Current business validation tools such as Middesk's Business Verification Dashboard ([Middesk Docs](#)) as shown in figure 2 present scores in a straightforward, accessible way, often accompanied by explanations to contextualize the result. These types of visualizations are critical for informing, engaging, and educating non-technical audiences. Dashboards are the primary visualization for communicating findings, particularly in business intelligence applications.

Figure 2. *Middesk's Dashboard Visualization*



Ivanković et al. (2021) developed seven "actionability" criteria for effective dashboard design:

1. Knowing the audience and their information needs
2. Managing the type, volume, and flow of information
3. Reporting data sources and methods clearly
4. Linking time trends to policy decisions
5. Providing data "close to home"
6. Breaking down the population into relevant subgroups
7. Using storytelling and visual cues

To apply these criteria effectively, the user, the data, and the insights must be aligned. Our dashboard design will integrate these principles to ensure that the scoring system is not only accurate but also meaningful and actionable for users.

A data-driven solution is only as strong as the data it's trained on. As the saying goes, "garbage in, garbage out"—flawed or biased input inevitably leads to flawed outcomes. Bias is rarely introduced intentionally, but it can seep into a model if left unchecked. Hort, Max, et al. describe three main strategies for bias mitigation: pre-processing, in-processing, and post-processing algorithms.

1. Pre-processing techniques occur before model training and involve adjusting the dataset to reduce inherent biases.
2. In-processing methods manipulate the algorithm itself, employing strategies like regularization, parameter tuning, or fairness-aware learning to improve outcomes.
3. Post-processing techniques occur after training and involve modifying outputs to correct for bias observed in predictions.

To effectively address industry and country of origin bias in our system, we will apply post-processing techniques, specifically, normalization by country maxima, sometimes referred to as min-max scaling (i.e., dividing each score by the maximum score within its country/industry and multiplying by 100), to supply the users with relative scores. Without intervention, companies from more established countries will consistently outscore businesses from other countries. However, the country of origin should not solely sway the legitimacy of a business. Therefore, bias mitigation is essential for ensuring that valid businesses from less established countries are evaluated fairly within our scoring system.

Data

To develop our global business confidence scoring algorithm, we draw from a rich and complex set of data sources provided by BrightQuery. These datasets span both U.S. and global businesses and include firmographic, financial, and digital presence indicators. Each source brings valuable signal to the table, though each also presents distinct challenges related to coverage, completeness, and integration. Below, we describe each source in depth, with attention to its relevance to our modeling goals and key data quality considerations.

Data Sources

1. U.S. Business Data Firmographics: This foundational dataset includes over 70 million U.S. companies and captures core business identity fields such as company name, location, industry classification (NAICS), company type (e.g., public/private), and LinkedIn URL.

2. Financials: The financials are split into four separate time-series tables: Most Recent Quarters (last four), Time Series – Annual, Time Series – Quarterly, and Time Series – Monthly. Together, these tables account for over 2 billion records, covering detailed financial metrics such as payroll, assets, EBITDA, operating expenses, and tax payments.

The two tables described above were used by BrightQuery to build their US-centric model. We used this table, specifically the business names and scores, to label the US businesses within the global dataset.

3. Global Data: The global data comes from Crunchbase, a platform that provides detailed information on over 3 million businesses across the globe. The provided data includes financial indicators such as operating status, revenue range, and IPO status. Furthermore, it includes online presence indicators, descriptions, location details, and more which are all essential business legitimacy indicators.

Challenges: Data completeness across the provided columns is uneven. Many columns have fill rates below 40%, and the columns that are above 40% are mostly categorical. This is due to the uneven amount of data provided by various countries. This unevenness directly influences our feature selection process as it is important to only consider features that provide enough information across all countries. Our focus is on high-coverage, high-variance, variables that show reliability across regions.

All datasets reside in Google BigQuery, and our team operates in a Google Cloud Platform sandbox environment optimized for analytical workflows. To stay within compute limits and maintain performance, we limit queries to subsets of 500,000 rows and pre-filter columns to retrieve only the most relevant fields. These constraints influence both our data preprocessing pipeline and the scalability of our modeling approach.

Descriptive Analysis

To assess the availability of key fields in the dataset, we analyzed their fill rates to better understand how these variables could contribute to our scoring model. Understanding the fill rates was essential for identifying which variables and datasets to include in our model. We aimed to select features that were well-populated and consistently available in both the global and U.S. data to ensure a standardized and reliable model. The dependent variable in our model is BrightQuery's U.S. confidence score, which ranges from 0 to 100—where 0 indicates no business legitimacy and 100 represents the highest level of legitimacy. Our goal is to predict this score using features that are also available for global businesses.

Figure 3.

Field coverage for subset of variables across all U.S. companies

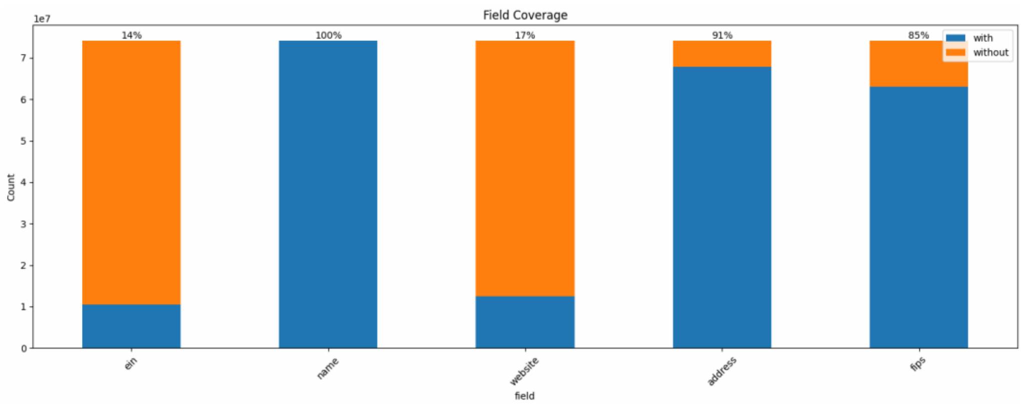


Figure 3 shows the fill rate of fields available in BrightQuery's firmographic data. This gave us an early indication that the presence of website and location identifiers were suitable variables for the model.

Figure 4.

Comparing the amount of fields in Crunchbase with above and below 40% coverage

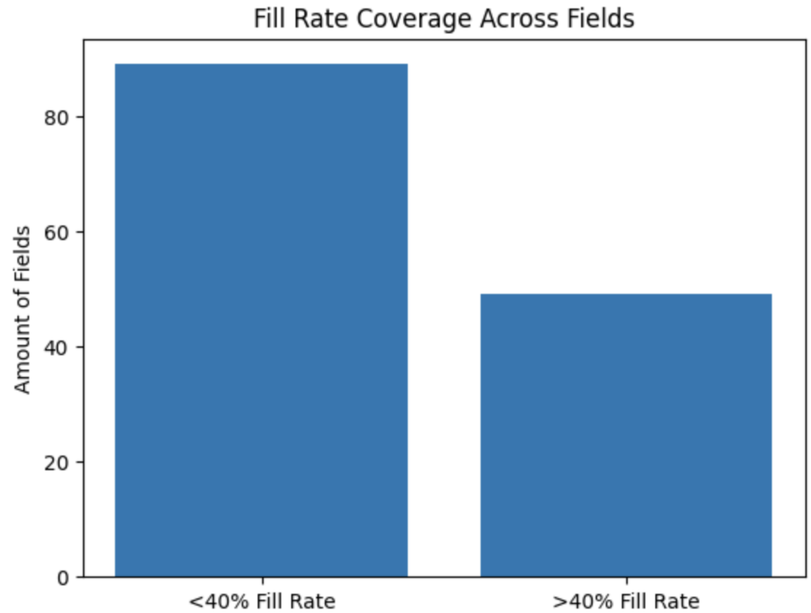


Figure 4 illustrates the fill rates among the provided fields. The majority of fields, 89 fields, have an insufficient fill rate of below 40%, indicating that we cannot use them in the model. 49 of the fields have sufficient coverage and can be considered for the model.

Figure 5.

Field coverage for a subset of high coverage variables across the Crunchbase dataset

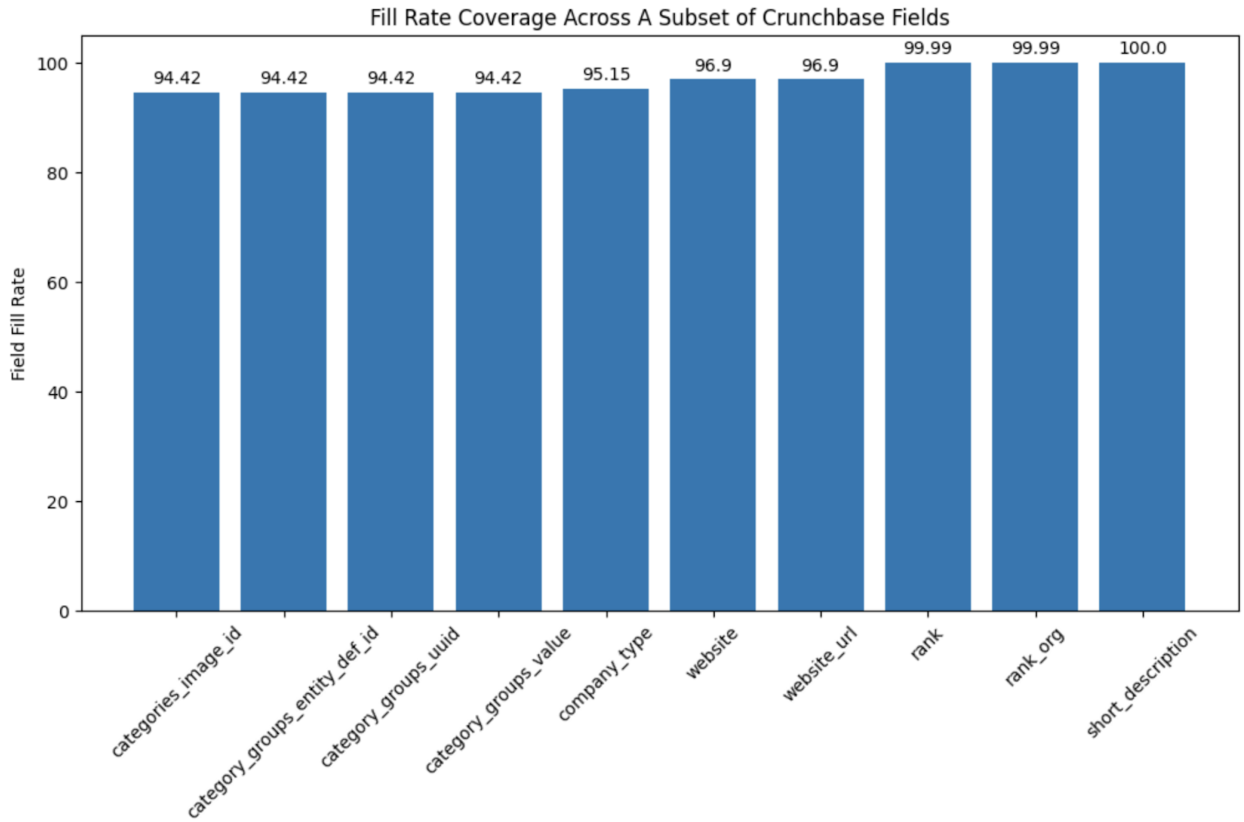


Figure 5 illustrates the coverage across a sample of high fill-rate columns within the Crunchbase dataset. The above fields, along with other high fill rate columns, were considered for the model.

Methodology

Analytical Plan

To develop a confidence scoring system that assesses business legitimacy, we focused on modeling U.S. business confidence scores using features available globally. After analyzing table join success rates and data completeness, we selected Crunchbase as the primary data source. This dataset not only includes essential online presence indicators but also covers both U.S. and global companies—critical for our modeling objectives.

We engineered a model that predicts the U.S. confidence scores generated by BrightQuery's U.S. model based on features that are available for global entities. Once the model reached satisfactory performance and interpretability, we applied it to score global businesses. These scores reflect absolute confidence and, to improve interpretability, we also computed relative scores—benchmarked within industries and across countries. This dual-scale approach enables clients to assess business legitimacy in both a global and local context.

Feature Engineering

We enhanced the dataset through a combination of derived features and thoughtful encoding strategies designed to reflect business legitimacy signals:

- **Company Age**: We transformed the founding date into a continuous feature representing the number of years since the company was founded. This provides insight into the longevity and stability of a business, which often correlates with legitimacy.

- **Data Freshness**: From the updated_at field, we calculated the number of days since the data was last updated. More recent updates are generally stronger indicators of an actively maintained and legitimate business profile.
- **Binary Presence Indicators**: For variables where the mere presence of a value carries more significance than the value itself, we applied binary encoding. This approach was used for fields like website_url, twitter, location_identifiers, phone_number, legal_name, and contact_email. For example, the existence of a website suggests legitimacy more strongly than the content or traffic associated with it. The same principle applies to having a Twitter account or a physical address.
- **Categorical Variables**: We applied categorical encoding to features where different levels signify different degrees of legitimacy. For instance, the number of employees and revenue range were both encoded into ordinal categories, as larger organizations tend to be more credible.

The status field was encoded to reflect stages of corporate lifecycle, with IPOs coded as 3, active operating companies as 2, acquired companies as 1, and closed businesses as 0. This captures the increasing order of perceived legitimacy. Similarly, the operating_status field was encoded as a binary indicator—1 for active and 0 for closed—emphasizing the distinction between ongoing and defunct businesses.

All feature transformations and encodings were reviewed and approved by the client prior to inclusion in the modeling pipeline.

Modeling Frameworks:

To model U.S. business confidence scores, we explored a variety of predictive frameworks. Our primary choice was Linear Regression, selected for its interpretability and ease of explanation. This model assumes a linear relationship between each input feature and the target score, meaning the final prediction is calculated as a weighted sum of the features plus a constant. Each weight, or coefficient, represents the strength and direction of that feature's impact on the confidence score. Linear regression also relies on several assumptions: independent observations, normally distributed errors, and constant variance of residuals across all levels of prediction. In other words, it assumes that the data points don't influence each other, that the errors in predictions are evenly spread out, and that the size of those errors stays consistent no matter what the prediction is. Because it provides clear insights into which factors most strongly influence perceived legitimacy, linear regression proved ideal for communicating results to business stakeholders.

To benchmark this model's performance, we also tested several alternatives, including Ridge Regression, Lasso Regression, XGBoost, Gradient Boosting, and Random Forest Regressor. These models can capture more complex, non-linear interactions between features and are generally more robust to outliers. However, they often sacrifice interpretability in exchange for accuracy. While linear regression remained our preferred model due to its clarity and practical usefulness, we used these additional models to assess whether their performance gains justified the added complexity. We evaluated all models using Root Mean Squared Error (RMSE)—an intuitive metric that reflects the average magnitude of prediction error, especially suited for our continuous response variable. Our goal was to achieve a low RMSE—specifically below 15—when compared to BrightQuery's U.S. confidence scores. The RMSE target of 15 was determined through discussions with the client, considering that the confidence scores range widely between 0 and 100. An RMSE of 15 was deemed a reasonable threshold, as it indicates predictions stay within the same general ballpark for assessing legitimacy despite the wide score range. Additionally, while accuracy was important, the primary objective was to maintain model interpretability, and an RMSE of 15 represented a balanced compromise between accuracy, simplicity, and interpretability. Setting a significantly lower RMSE goal might have led to a trade-off where we sacrificed interpretability for better model performance. However, we didn't want to make that compromise because interpretability was far more crucial. Given that the model's accuracy was already sufficient, it was essential for the client to understand the key drivers behind the legitimacy score. This transparency is important for providing rationale and enabling users to further explore the reasons behind the score. Since global businesses lacked pre-existing scores, we applied the trained model to generate confidence scores for businesses worldwide. To validate these global scores, we performed intuitive checks on a sample of 10 well-known companies—such as Amazon, Microsoft, and Apple—verifying that their scores were relatively higher compared to others. This approach helped confirm that the model's outputs aligned with expectations, reflecting the established credibility and legitimacy of these prominent companies.

Findings

We filtered the dataset to only include variables with a fill rate above 40% to ensure reliable coverage. Presence-based fields such as website URL, Twitter, Facebook, location ID, phone number, and contact email were encoded as binary indicators. Revenue range, employee count, and lifecycle fields (active, IPO, acquired, closed) were encoded to capture operational scale and status in a way models could interpret.

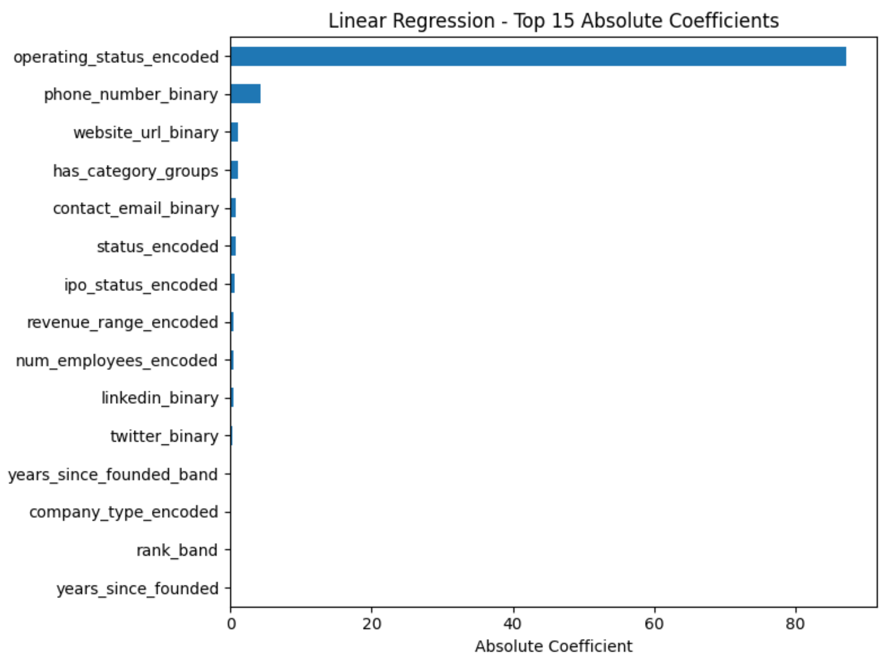
We trained and evaluated Linear Regression, Ridge, Lasso, Random Forest, Gradient Boosting, and XGBoost to predict the BrightQuery confidence score.

Model	MAE	RMSE
Linear Regression	10.37	12.48
Ridge Regression	10.37	12.48
Lasso Regression	10.46	12.51
Gradient Boosting	10.18	12.29
XGBoost Regression	10.13	12.27
Random Forest	10.88	13.48

Although Gradient Boosting and XGBoost achieved slightly better RMSE scores, Linear Regression was selected as the primary model because its coefficients are transparent and easy to explain, making the results defensible and easier to maintain in operational use.

Figure 6.

Feature Importances in Linear Regression Model for Confidence Scoring



Across all models, the most impactful feature was operating_status_encoded, which distinguishes active from inactive companies. phone_number_binary followed as a strong indicator of legitimacy, with website URL, status, and contact email also contributing meaningfully. Other features such as revenue range and employee tiers had smaller but still positive effects. Country context also proved important, allowing confidence scores to be interpreted fairly within each region.

Additionally, to validate our model, we conducted intuitive checks to assess the reasonableness of the generated scores for select companies. For instance, Amazon received a score of 92, NVIDIA scored 91, and Microsoft scored 92. These results aligned with expectations, as these are well-established, credible global companies. Observing such scores for high-profile examples gave us added confidence in the reliability and consistency of the global scoring approach. Our highest-performing countries were the United States, United Kingdom, India, Germany, and Canada. In terms of industry, the top scorers came from manufacturing, government and military, design, consumer goods, and administrative services. This alignment made intuitive sense because these sectors and regions often emphasize structured processes, rigorous standards, and innovation—factors that likely contributed to their strong performance.

Discussion

For our model, one of the objectives was to meet an accuracy target of RMSE below 15. The model we chose, Linear Regression, was able to meet this target with an RMSE of 12.48. We also needed the model to be easily interpretable, which a Linear Regression provides, with coefficients representing the strength and direction of that feature's impact on the confidence score. Other models, such as Gradient Boosting, did not offer direct feature weights for the scoring process, which was essential for interpretability—which is why they were not selected. Although linear regression did not achieve the lowest RMSE among all models, it was chosen for its interpretability and its ability to provide clear, quantifiable weights that explain how scores are generated. This enables the client to identify which features were most influential in determining the quantitative metric for a business's legitimacy. These insights can help justify pursuing specific partnerships or highlight areas that warrant further investigation for deeper understanding. We chose Crunchbase as our data source because it provided consistent fields for both U.S. and global businesses, allowing us to use the same features across regions. This enabled the development of a standardized model for scoring business legitimacy worldwide.

Our analysis found that operating status and the presence of a phone number are strong contributors to higher confidence scores for U.S. businesses in our model. This suggests that these two factors are key indicators of business legitimacy in global data. We also observed that the presence of a contact email and a website URL positively influenced confidence scores, highlighting the importance of a verifiable online presence. These findings provide valuable insights for clients. When a business is actively operating and offers clear contact details—especially a working phone number and online touchpoints like email and a website—it signals a higher likelihood of legitimacy. Clients can use these attributes as reliable indicators when assessing the credibility of businesses, particularly in unfamiliar regions or large datasets. This enables more accurate risk assessments, faster onboarding processes, and better-informed business decisions.

One limitation of our approach was difficulty linking fields across global datasets, which prevented us from creating a fully cohesive dataset. As a result, one limitation of our score is that it does not incorporate deep financial insights or financial ratios, which could enhance the accuracy of the legitimacy metric. This presents an opportunity for improvement in future iterations. We also successfully created relative scores while minimizing bias by first generating absolute scores through our model, then normalizing them by country and industry. This approach allows businesses to be compared on a consistent scale within their specific geographic or industry context, rather than globally. As a result, users can more effectively evaluate and investigate businesses within targeted countries or sectors. Finally, we developed a dashboard with filtering capabilities that allows users to view scores, examine the feature importance behind each score to understand the key drivers, and compare business scores side by side.

Conclusion

Our final deliverable includes our linear model validated on existing confidence scores for U.S. businesses within the global data, as well as an interactive dashboard. The interactive dashboard allows users to search for specific businesses to view confidence scores, features that are the key drivers behind each score, and a relative score based on country. Users are also able to directly compare two businesses. An exploration of trends across industries and countries is also possible, as the dashboard allows for filtering scores by country, industry, or both, as well as displaying the highest-scoring businesses by country and industry. By visualizing these factors, users can gain deeper insights into a company's credibility.

For BrightQuery, our model's global capabilities allow for an expansion of their product offerings once refined and put into production. Because the offering is new and unprecedented in its scope, it is difficult to accurately estimate impact in terms of monetary value. BrightQuery has estimated that there will be around 3,000 initial target companies for the new product, with revenue being generated from two points- origination and monitoring. Origination being the initial check run on a company by a client, and monitoring consisting of periodic checks run by the client after the initial lookup.

In the future, steps can be taken to further improve the capabilities of the model. This can be done through improvements in the data available to be used in the model, primarily through building effective methods to join the global datasets. This would allow access to features not available in the Crunchbase data, such as detailed location information. Changes and refinements to the model may also occur due to BrightQuery updating their U.S. confidence scoring model. The dashboard interface for the model will also likely evolve in response to client demands.

References

Davron. "*History of Professional Licensure.*" Davron, 2 Jan. 2020, https://www.davron.net/the-history-of-professional-licensure/.

Dwyer, D. W., Chen, A., Kocagil, A. E., & Tian, P. (2004). *Moody's KMV RiskCalc™ v3.1 model: Next-generation technology for predicting private firm credit risk*. Moody's KMV Company. https://www.moodys.com/sites/products/ProductAttachments/RiskCalc%203.1%20Whitepaper.pdf

Greif, Avner. "*Chapter 4: Securing Property Rights from the Grabbing Hand of the State: The Merchant Guild.*" Institutions and the Path to the Modern Economy, n.d., https://web.stanford.edu/~avner/Greif_Instutions/0%204%20Chapter%204%20Guild1.pdf.

Hort, M., Zeng, Y., Malik, K., & Weller, A. (2024). Bias mitigation for machine learning classifiers: A comprehensive survey. *ACM Journal on Responsible Computing, 1*(2), 1–52. https://doi.org/10.1145/3631326

Ivanković, D., Barbazza, E., Bos, V., Brito Fernandes, Ó., Gilmore, K. J., Jansen, T., Kara, P., Larrain, N., Lu, S., Meza-Torres, B., Mulyanto, J., Poldrugovac, M., Rotar, A., Wang, S., Willmington, C., Yang, Y., Yelgezekova, Z., Allin, S., Klazinga, N., & Kringos, D. (2021). Features constituting actionable COVID-19 dashboards: Descriptive assessment and expert appraisal of 158 public web-based COVID-19 dashboards. *Journal of Medical Internet Research, 23*(2), e25682. https://www.jmir.org/2021/2/e25682

Middesk. (n.d.). *Business verification tool dashboard overview*. Retrieved April 29, 2025, from https://docs.middesk.com/docs/overview-

Moody's Investors Service. (2021, November 12). *Business and consumer services*.
https://ratings.moodys.com/api/rmc-documents/356424