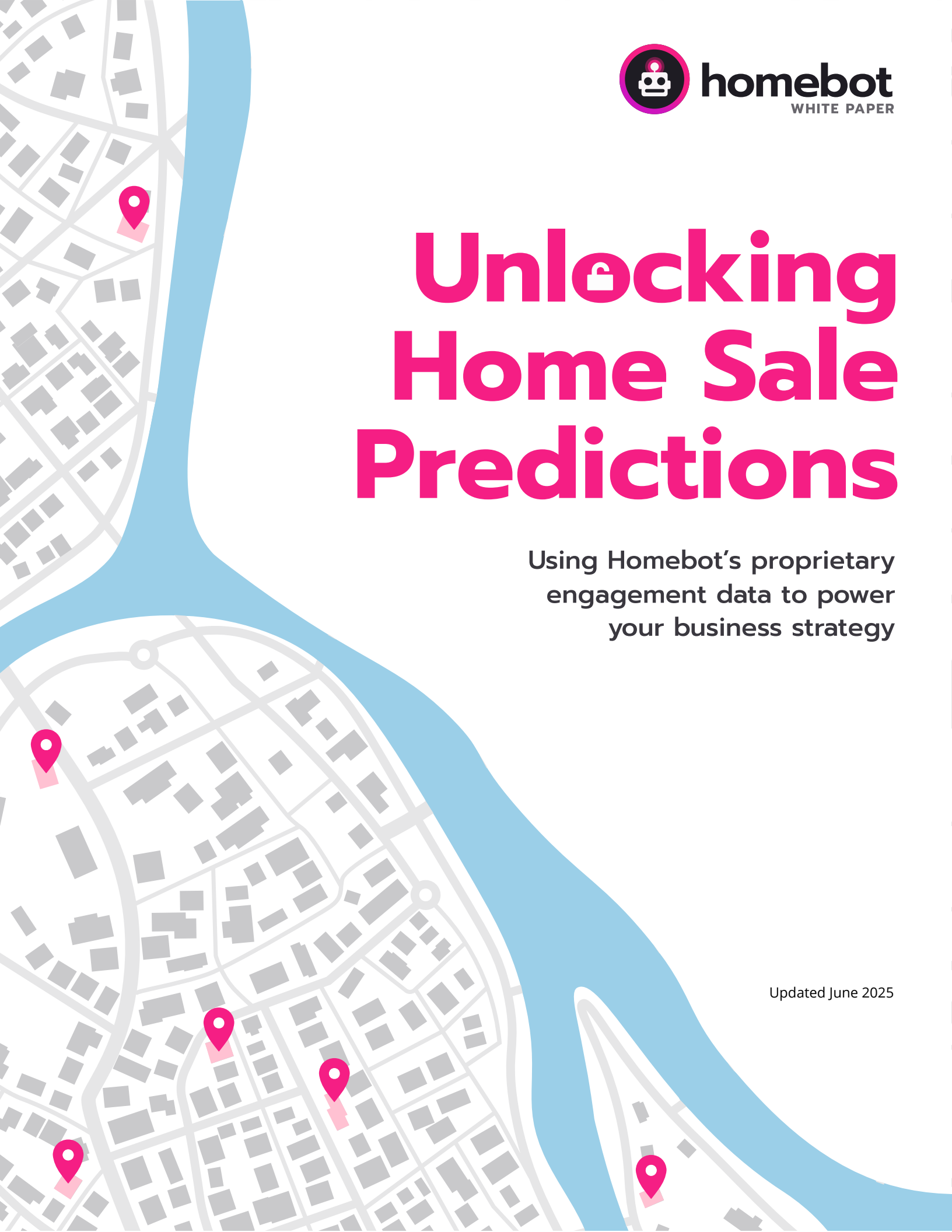
homebot
WHITE PAPER

Unlocking Home Sale Predictions

Using Homebot's proprietary engagement data to power your business strategy



Updated June 2025

Introduction & Findings

The Homebot Likelihood to Sell Score is an innovative machine learning (ML) model designed to accurately predict your homeowners' likelihood to list and sell their home within the next 9 months.

By leveraging our proprietary engagement data and home & loan information, we offer **unprecedented insights into homeowners' selling intentions**.

304M
ROWS OF DATA ANALYZED

70%
OF ACTUAL MOVES CORRECTLY PREDICTED IN TOP 30% OF SCORES

The Prediction Model

We leveraged Homebot's engagement plus home & loan data to create a comprehensive view of potential sellers

Engagement

+ CMA REQUESTS
+ HOMEBOT VIEWING FREQUENCY
+ LISTING VIEWING FREQUENCY
+ MESSAGE ACTIVITY
+ ADDITIONAL TOUCH POINTS

Home & Loan

+ LOAN RATE
+ MONTHS IN HOME
+ ZIP CODE
+ EQUITY POSITION

More details on page 7 →

Results

We analyzed how well our predicted probabilities matched up with actual sellers

70%

of actual moves were correctly predicted with **likely to sell scores**

5 Sellers in a sample of 100 **random** clients

VS USING HOMEBOT SCORES:

30 Sellers in a sample of 100 **likely to sell** clients

43 Sellers in a sample of 100 **highly likely to sell** clients

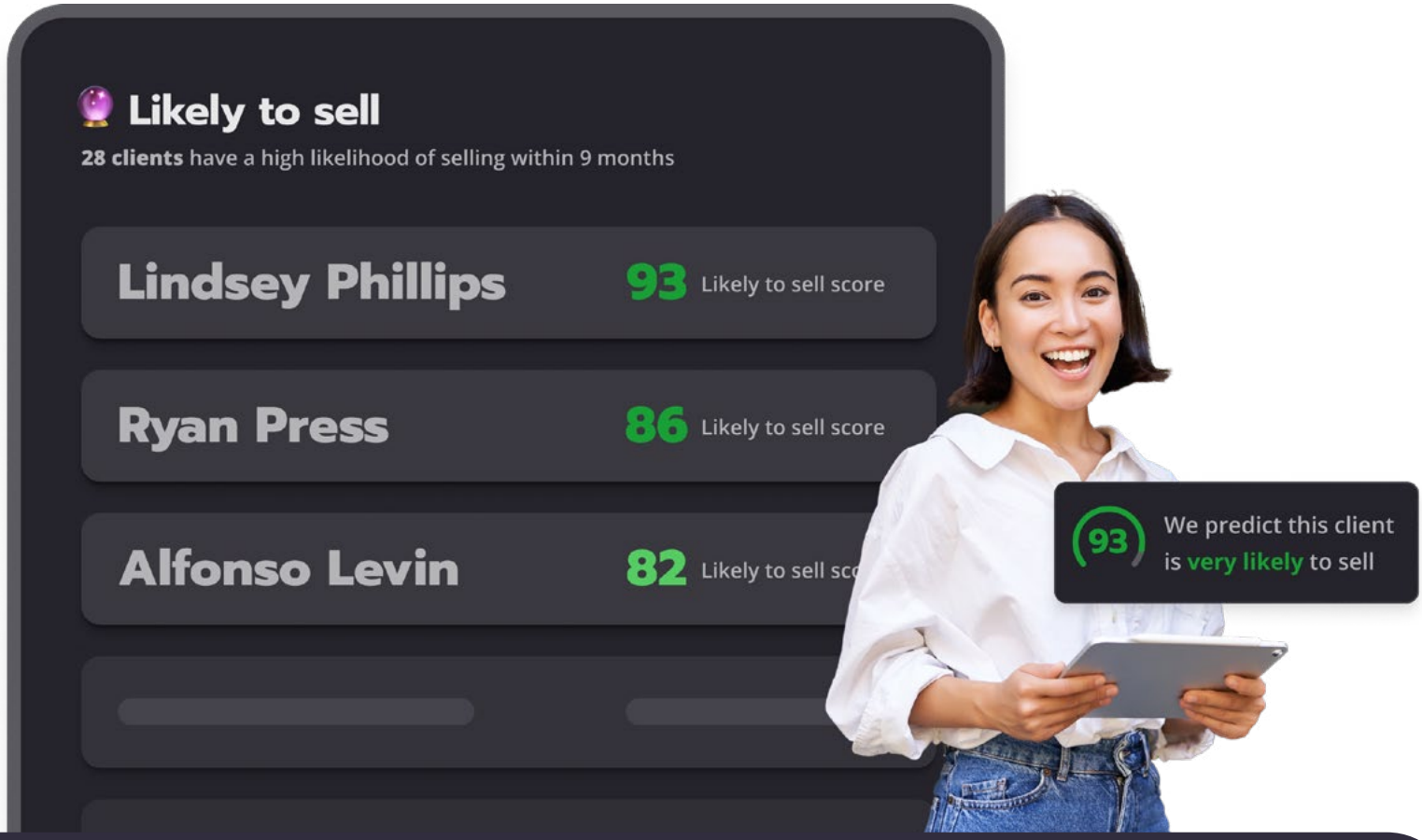
This results in an **8x** improvement in sales and marketing efficiency

Using the scores

1. Save 70% on marketing by targeting only likely-to-sell clients
2. Pre-curate a listing for a client's home to help win their business
3. Predict pre-payments and focus your recapture efforts

Accessing the scores

Likelihood to sell scores are available inside the Client Engagement Portal on the client details page as well as via a Key Client List that highlights the clients that are likely to sell.



Easy exporting

Enterprise customers can request a CSV of all likelihood to sell scores for their clients and even receive an export daily

Conclusion

By leveraging the unique insights offered by Homebot's platform, lenders and real estate agents **gain a significant advantage** in proactively engaging with potential sellers, optimizing marketing campaigns, and ultimately **driving business growth.**

As Homebot continues to innovate and enhance its engagement tools, the predictive power and impact of this model will continue to expand, helping you get in front of leads even faster.

Key Takeaways

- **CURATED LIKELY TO SELL LISTS AND CLIENT SCORES**
- **SCORE AND CONTACT LISTS EXPORTED WEEKLY**
- **MORE PREDICTIVE SCORES AND MODELS COMING SOON**

Ready to get in front of your sellers first?

Reach out to see how you can start leveraging seller prediction scores!

GET IN TOUCH

Or shoot us an email: **sales@homebot.ai**

Continue for more details on the data model and performance →

Data Quantity & Quality: The Foundation for Accurate Predictions

The efficacy of any machine learning model hinges upon the quantity and quality of the data used for training. For our home sale prediction model, we have prioritized the collection and curation of a robust dataset that ensures both **representativeness** and **reliability**.

Data Quantity

Our model benefits from a vast dataset encompassing millions of homeowner records accumulated over a period of four years (January 2020 - January 2024).

In total, we leveraged over 294M rows of data for training and testing the model. This extensive data collection allows the model to learn complex patterns and relationships within homeowner behavior and market dynamics, leading to more accurate and generalizable predictions.

Data Quality

Engagement Data Accuracy: Homebot's platform meticulously tracks and timestamps every user interaction, ensuring the accuracy and precision of our engagement data. This granular data collection provides a reliable foundation for understanding homeowner behavior and identifying meaningful trends.

Home & Loan Data Validation: We employ rigorous data validation processes to ensure the accuracy and completeness of home & loan information, such as loan details and property addresses. This includes cross-referencing with public records and implementing data cleansing techniques to address any inconsistencies or missing values.

Feature Engineering: The model incorporates carefully engineered features designed to capture the most relevant aspects of homeowner behavior and market conditions. This involves thoughtful selection of engagement metrics, home & loan variables, and calculated features such as the loan rate difference to benchmark.

Data Cleaning and Preprocessing: We implement data cleaning procedures to address missing values, outliers, and inconsistencies within the dataset. Numerical and categorical variables are standardized and formatted to prepare for model training.

IMPACT ON MODEL PERFORMANCE

The quantity and quality of our data directly contribute to the model's exceptional performance. The vast dataset allows for comprehensive learning, while the emphasis on data accuracy and feature engineering ensures that the model captures meaningful relationships and avoids biases. This ultimately leads to more reliable and actionable predictions, empowering lenders and real estate agents to make informed decisions and optimize their engagement strategies.

Home & Loan Details

LOAN RATE DIFFERENCE TO BENCHMARK
This feature serves as a proxy for the homeowner's potential financial gain from selling their current home and securing a new mortgage at a lower interest rate. A larger difference between their existing rate and the current market benchmark suggests a stronger incentive to sell.

MONTHS IN HOME
This feature reflects the duration of homeownership, which is historically linked to the propensity to sell. The model captures the natural lifecycle of homeownership and its influence on selling behavior.

ZIP CODE (FIRST 2 DIGITS)
Utilizing the first two digits of the zip code provides regional information that reflects broader economic and market trends without being overly specific. This balances the need for granularity with the requirement of sufficient data for reliable predictions.

EQUITY POSITION
Homeowners' equity position reflects their potential financial flexibility. Higher equity suggests a greater ability and incentive to make significant housing decisions such as selling, tapping into their equity, or investing in new properties.

Engagement Data Details

CMA REQUESTS
A homeowner's request for a CMA is a strong indicator of selling intent, as it suggests an active interest in understanding their home's current market value and exploring the potential for a sale.

HOMEBOT VIEWING FREQUENCY
Frequent interaction with the Homebot platform signifies that a homeowner is actively engaged with their home finances and the real estate market. The number of days they've accessed Homebot within a specific timeframe serves as a valuable indicator of potential selling intent.

LISTING VIEWING FREQUENCY
The extent to which a homeowner explores listings on the Homebot platform reflects their interest in the current market and potential consideration of alternative properties.

MESSAGE ACTIVITY
Communication with lenders or real estate agents through Homebot suggests a proactive approach to exploring financial options and potentially preparing for a sale.

ADDITIONAL ENGAGEMENT FEATURES
The model incorporates a variety of additional engagement metrics, such as the use of financial calculators and interaction with educational resources, to paint a comprehensive picture of homeowner behavior and intent.

Model Performance and Evaluation: Quantifying Predictive Power

To assess the effectiveness of our home sale prediction model, we employed a comprehensive evaluation strategy encompassing various performance metrics and in-depth analysis.

Metrics and Results

Area Under the ROC Curve (AUC): The model achieved an impressive AUC of **0.78,** as visualized below. This indicates a strong ability to distinguish between homeowners likely to sell and those likely to stay put.

Balanced Accuracy: Considering the imbalanced nature of the data, with a significantly larger proportion of non-sellers, we utilized balanced accuracy to ensure a fair evaluation. The model attained a balanced accuracy score of **0.70,** demonstrating its ability to effectively predict both positive and negative cases.

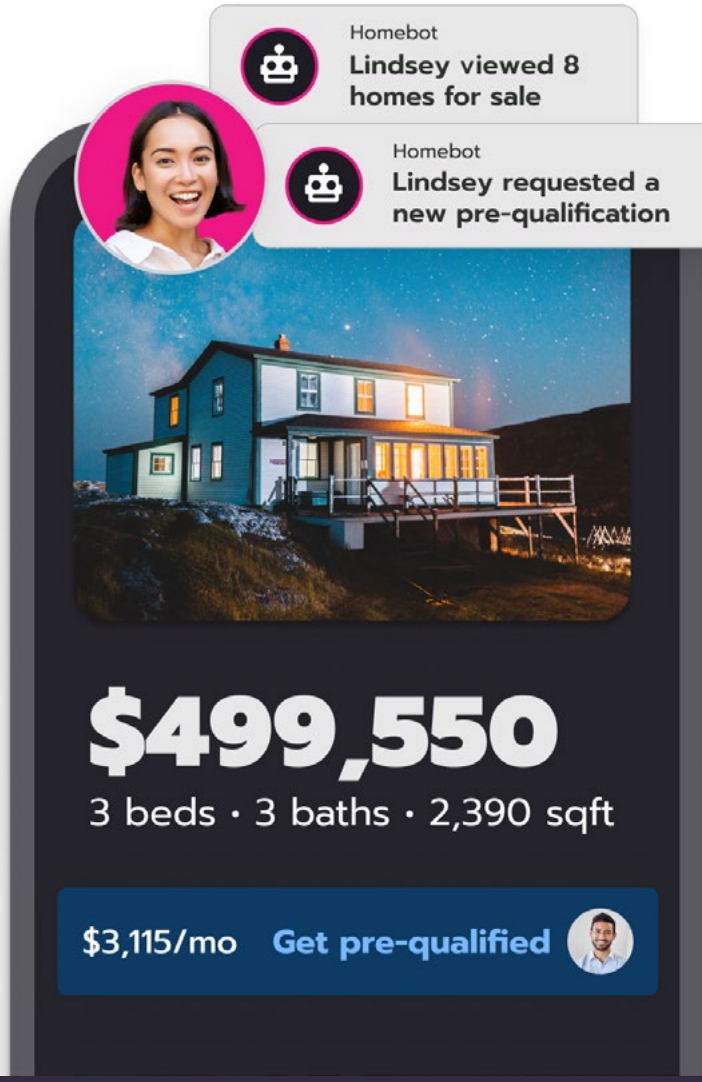
True positive rate: To emphasize the importance of correctly identifying potential sellers, we also calculated the True Positive Rate (TPR), which measures the number of actual sellers the model correctly identified. The TPR of the model is 69.9%, meaning the model correctly identified 69.9% of actual sellers correctly.

Receiver Operating Characteristic Curve

About Homebot

Homebot is a **Homeownership Platform** surfacing actionable data first so you can capture the client before the competition.

With Homebot, we tell you which clients are most likely to need a loan and when. This means you get to these active leads first. Our platform identifies origination opportunities so that you focus on capturing transactions instead of chasing clients who aren't ready.



FOR MORE INFORMATION, VISIT:

homebot.ai Homebot LinkedIn

homebot

Ready to learn more?

SCHEDULE A DEMO