Swarmalytics

MODELING WITH ARTIFICIAL INTELLIGENCE TO PREDICT KEY METRICS FOR NEW HOMES: THREE USE CASES

Predictive models based in AI can predict the listing price, sale price, and time on market for a new home with little margin of error, decreasing uncertainty on the part of home builders and buyers alike.

Swarmalytics

Summary

Swarmalytics Inc. has developed a foundational set of models that use artificial swarm intelligence to predict the initial listing price, time on market, and final sale price of newly constructed homes in the United States. These models aim to replace inefficient home pricing methods used by home builders that do not consider market behavior or trends. By combining AI technology with a wide variety of data sources, Swarmalytics has developed models that take into account factors that affect the real estate market but which go unnoticed or overlooked by analysts and traditional methods. These models will allow home builders to price their homes more accurately, allocate resources more effectively, and adjust their sales strategies as circumstances dictate. The use of artificial swarm intelligence represents a significant advancement in the field of property technology, allowing builders to make more informed decisions and ultimately generate higher profits.

Introduction

The home construction industry in the United States is a major contributor to the nation's economy. However, the industry is plagued by inefficient pricing methods that do not account for market behavior or trends. This results in twin sources of lost revenue; homes that are underpriced during a hot market leave money on the table while homes that are overpriced during a cold market remain unsold for long periods of time, incurring maintenance costs on the part of the builder. To address this problem, Swarmalytics has developed predictive models using artificial swarm intelligence to predict the initial listing price, time on market, and final sale price of newly constructed homes in the United States. The goal of these models is to provide home builders with tools needed to plan for the future and generate higher profits.

The predictive models that Swarmalytics has developed can serve as a strong foundation for future pricing strategies by US home builders. While the results shown later in this document are encouraging, future models that incorporate data owned by specific home builders have the potential to greatly surpass these established benchmarks. Builder-specific models can learn from the geography, typical customer profile, and market behavior of a particular home builder to create predictions that are more accurate and more actionable.

Background

The New Home Market in 2023

The United States housing market has experienced tremendous growth in recent years, with nearly two million new single-family houses sold in the first quarter of 2023[1]. However, despite this growth, homebuilders face significant challenges when it comes to pricing their newly constructed homes. Home prices rose steadily from 2020 through 2022, reaching a median value of about $480,000 by Q4 2022 before falling sharply in Q1 2023[2]. Even marginal inefficiencies in pricing can result in billions of dollars in lost revenue annually.

Currently, prominent pricing techniques for home builders involve either cost-plus pricing or comp-based pricing. Cost-plus pricing is reliant solely on factors related to construction, such as material costs, number of rooms, lot price, and on-site amenities. Comp-based pricing is a slightly more sophisticated approach that includes additional data points such as local amenities. However, both techniques are unreliable in calculating the optimal price for a home, as they only consider a limited number of data points.

To improve pricing accuracy, homebuilders need to consider a wider range of market factors, such as crime statistics, traffic flows, or the performance of local schools. However, gathering and utilizing this

data is a complex task. The real estate world offers vast amounts of data that can be used for analysis, including housing data, population characteristics, and regional and local economic data.

Advanced data collection and manipulation techniques, developed by Swarmalytics Labs, have overcome the challenge of accumulating, and utilizing this data. The company has created a massive proprietary data set for analysis, using advanced machine learning and swarm intelligence algorithms to analyze large amounts of data. The company's predictive modeling approach has significant potential to reduce uncertainty and optimize home pricing for homebuilders, providing a new solution for pricing newly constructed homes.

Artificial Swarm Intelligence

Manual analysis of this data is out of the question; data of this magnitude can only be efficiently analyzed using artificial intelligence to search for predictive relationships at extremely high speeds. Swarmalytics' artificial swarm intelligence engine has this capability. Research into swarm intelligence and its applications dates back to the late 1980s. Examples of early milestones include the creation of Boids, an artificial life program written by Craig Reynolds in 1986 using simple rules to create a simulated flock of birds[3], and the 1992 publication of Marco Dorigo's Optimization, Learning, and Natural Algorithms, which laid the groundwork for ant colony optimization algorithms[4]. In the decades since, the study of swarm intelligence has continued to bear fruit. For example, Southwest Airlines' efficient open-seating policy was born out of swarm intelligence, which is also used to navigate their pilots to airport gates most efficiently[5]. More recently, Fortune Magazine explored the power of swarm intelligence to transform the financial world in its November 2021 article Following the Flock – The Science of Meme Stocks and Manias[6], and a May 2021 article from Harvard Business Review highlighted the innovative uses of swarm intelligence across companies such as Hewlett-Packard, Unilever, and Capital One[7].

Model Development

Data Sources

A database containing every US residential property from First American Real Estate Solutions served as the foundation of the modeled data. Swarmalytics then appended a wide variety of additional data sources, including proprietary data sources, in order to develop the predictive models. Data from multiple listing services was used to obtain additional information about each home. Additional data sources included the Internal Revenue Service as well as the Securities and Exchange Commission, providing household incomes, business activity, and other economic characteristics of the home's surrounding area. Geographic proximity data containing the distance of the home from schools, hospitals, parks and other amenities, as well as the distance of the home from businesses of various types was also included. All of these additional data sets and more were appended to the modeled data, enhancing it with additional independent variables to weigh as potential factors in each model's equation.

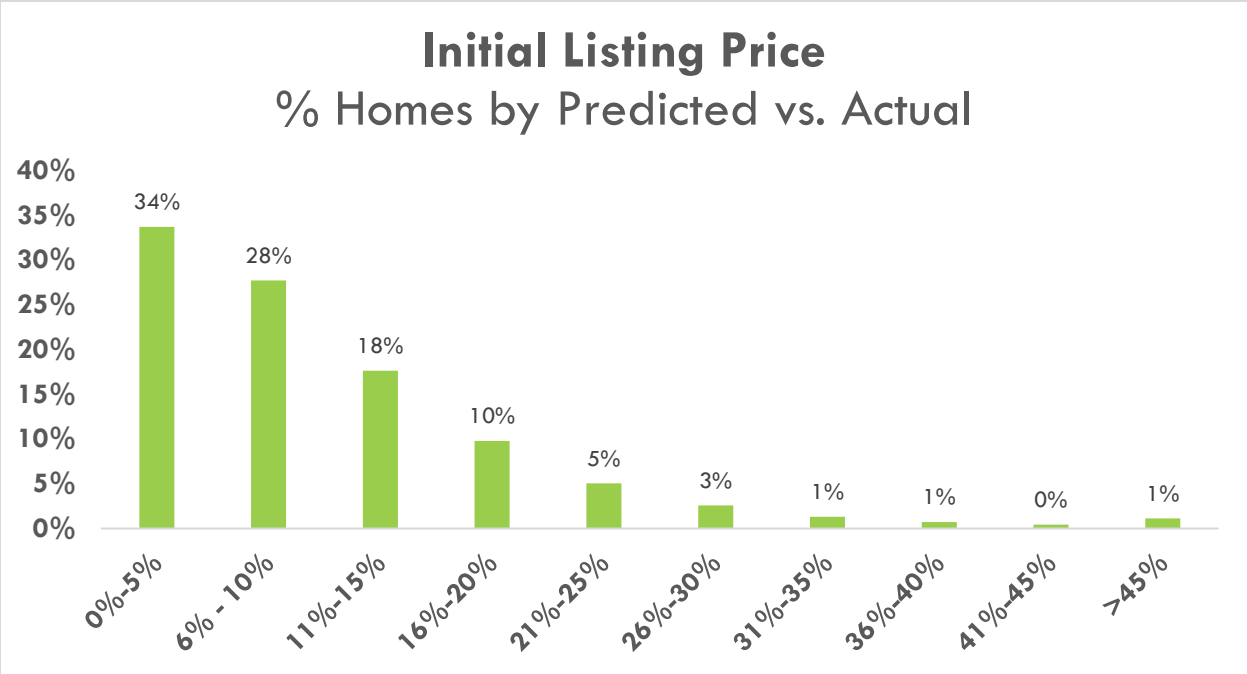
Modeling Approach

Swarmalytics' predictive models were developed using artificial swarm intelligence to train models on huge data sets of newly constructed US homes. This approach involves separating problems into homogeneous subsets, using massive parallel processing to find solutions to each problem, and finally combining the solutions to these problem subsets into a comprehensive whole. The solution to each problem subset is formed using machine learning and genetic algorithm techniques, wherein candidate equations are generated and then compete against one another to create more accurate predictions.

Predictive Models

Initial Listing Price Model

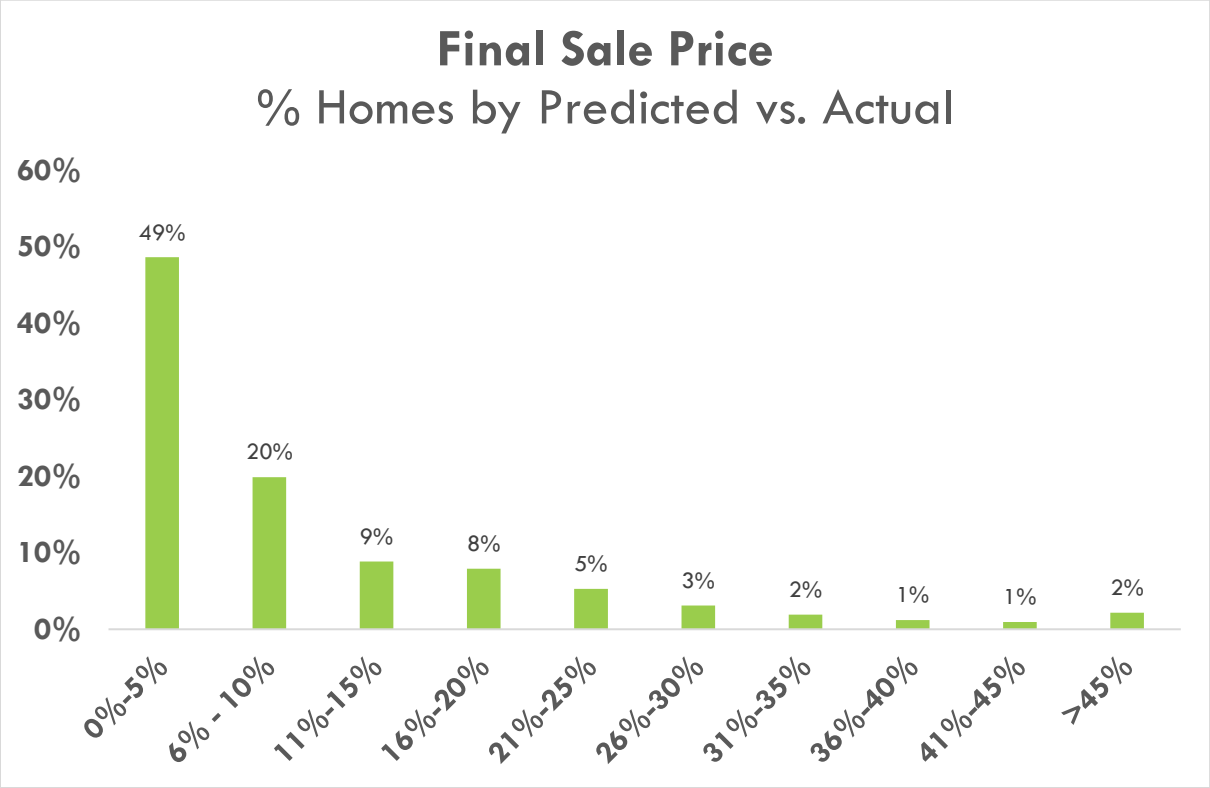
The initial listing price model was based on a data set of 24,999 newly constructed homes across the United States with the goal of predicting the listing price of a home when it was first put on the market. Changes in this listing price during the course of the home's time on market, such as a home's price being lowered to attract more potential buyers, were not considered during this analysis. The final model equation was driven primarily by attributes of the home such as square footage, as well as listing prices of comparable homes. Community attributes, census demographics, and local land and business variables also played a part in predicting initial listing price. The results of this model are promising; 34% of predicted listing prices fell within 5% of the actual listing price, while 62% of predictions were within 10% of the actual listing price.



Swarmalytics

Final Sale Price Model

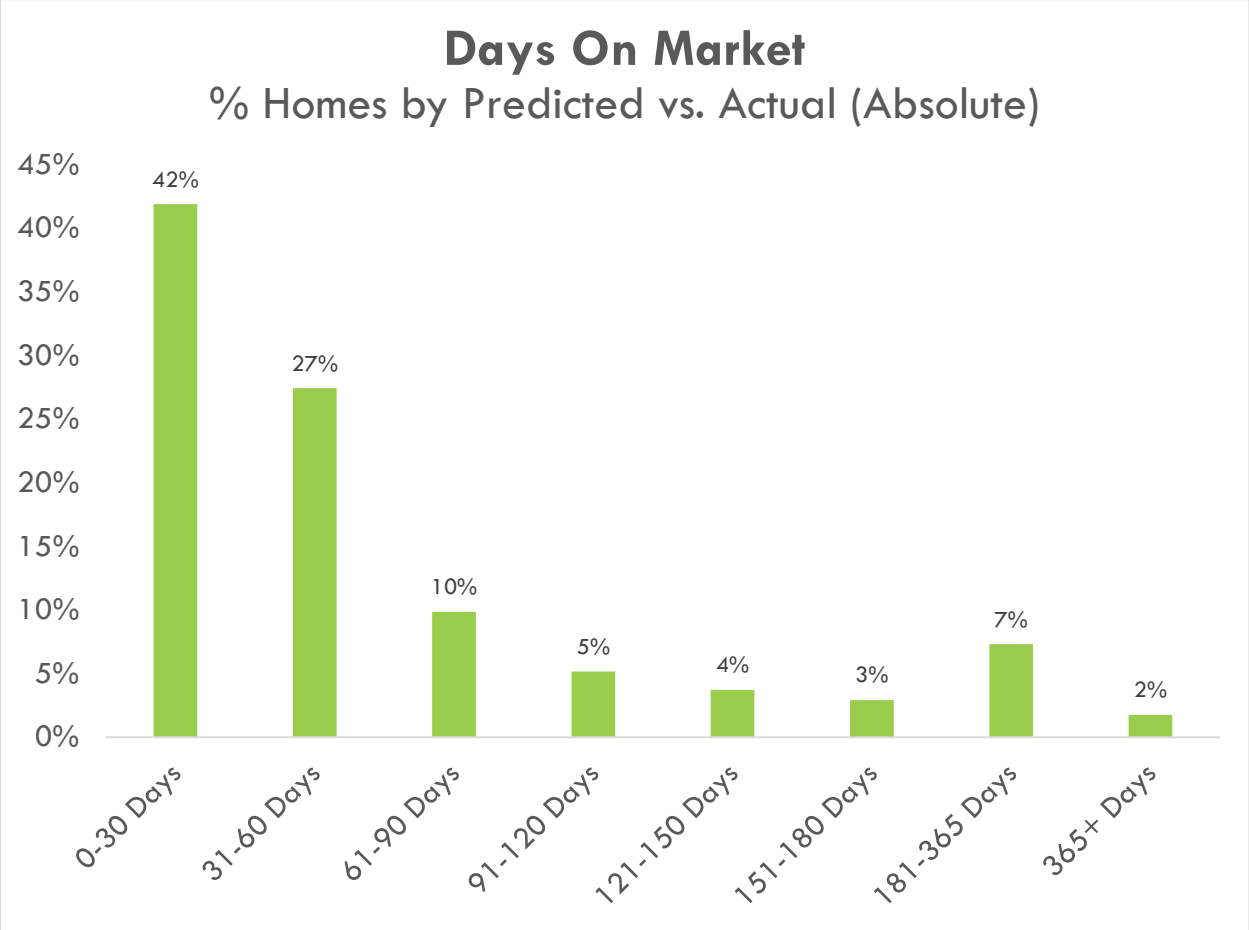
The final sale price model used a data set of 417 newly constructed and recently sold homes in the US as input values. Like the initial listing price model, the model equation was driven primarily by home attributes and sale prices of comparable homes, with community attributes such as the presence of tennis courts also factoring into the model. The results of this model appear even stronger than the initial listing price model. Among the predicted sale prices, nearly half were within 5% of the actual sale price. In total, nearly 70% of predictions were within 10% of the actual value.



Swarmalytics

Time on Market Model

The time on market model sampled a snapshot of 94,233 listed homes across the country. The dependent variable for this model was the total number of days between a home being listed for sale and that home being sold. The final equation for the time on market model was less reliant on home attributes, focusing instead of data elements related to the builder of the home such as their homes' median time on market. 40% of predicted time-on-market values were within 30 days of the actual value.



Refining the Models

The models developed by Swarmalytics can be further refined in a variety of ways. Building separate models for different geographic regions of the US may improve predictive performance. Swarmalytics also plans to perform meta-analysis to assess the degree to which model accuracy varies across segments of the housing market. It is possible that the predictions generated by Swarmalytics' current models are less accurate for homes with particular characteristics such as size, urban vs rural location, price, and so forth. As these weak points are discovered, specialized models can be trained to predict listing price, sale price, and time on market more accurately for these particular segments of the housing market. Builder-specific models can also be constructed, utilizing data specific to that builder for greater predictive power.

Conclusion

Swarmalytics has developed models using artificial swarm intelligence to predict the initial listing price, time on market, and final sale price of newly constructed homes in the United States. These models seek to eliminate inefficient home pricing methods currently used by home builders, which do not account for market behavior or trends. Swarmalytics' approach to predicting home prices using artificial swarm intelligence has several benefits over traditional methods used by home builders. The use of AI technology in these models allows for the processing of vast amounts of data and the identification of complex patterns. Additionally, the use of multiple data sources, such as IRS data and geographic proximity data, enables a more comprehensive understanding of the housing market and its underlying dynamics.

By predicting the initial listing price, time on market, and final sale price, Swarmalytics' models allow home builders to price their homes more accurately, reducing the risk of underpricing and overpricing during hot and cold markets, respectively. The use of AI in predicting home prices also allows builders to plan for the future and generate higher profits. By accurately predicting the final sale price of a home, builders can make better decisions about which projects to pursue and how to allocate resources. For example, if the predicted sale price of a home is significantly higher than the cost of construction, the builder may choose to invest more in high-end finishes or amenities to increase the home's value and ultimately, the sale price. Furthermore, the ability to predict the time on market can help builders adjust their marketing and sales strategies accordingly. If a model predicts a long time on the market, for example, the builder may decide to adjust the price or marketing tactics to attract more potential buyers and reduce the time on the market.

In summary, the use of artificial swarm intelligence by Swarmalytics to predict the initial listing price, time on market, and final sale price of newly constructed homes in the United States represents a significant advancement in the field of property technology. By integrating multiple data sources and using AI technology to identify complex patterns in the data, these models provide home builders with a more accurate and comprehensive understanding of the housing market, allowing them to make more informed decisions and ultimately, generate higher profits. The results of these models show great promise but should also be seen as the foundation for future models that incorporate builder-specific data. Future work by Swarmalytics will also further refine these models and expand their application to different segments of the housing market.

About Swarmalytics

Swarmalytics was founded in 2021 by CEO Doug Newell and CTO Dan Koehler with the goal of using artificial swarm intelligence to power insights and create value. To learn more about Swarmalytics, please visit www.swarmalytics.com.

Citations

1. U.S. Census Bureau and U.S. Department of Housing and Urban Development, New One Family Houses Sold: United States [HSN1F], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/HSN1F, February 17, 2023.
2. U.S. Census Bureau and U.S. Department of Housing and Urban Development, Median Sales Price of Houses Sold for the United States [MSPUS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/MSPUS, February 17, 2023.
3. Reynolds, Craig W. (1987). "Flocks, herds, and schools: A distributed behavioral model". Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'87). ACM. 21 (4): 25–34.
4. M. Dorigo, 1992. Optimization, Learning and Natural Algorithms, PhD thesis, Politecnico di Milano, Italy.
5. "What Can Ants Teach Us?" CBS News, www.cbsnews.com/news/what-can-ants-teach-us/. Accessed 16 Nov. 2022.
6. "Planes, Trains and Ant Hills." Science Daily, 1 Apr. 2008, web.archive.org/web/20101124132227/www.sciencedaily.com/videos/2008/0406-planes_trains_and_ant_hills.htm. Accessed 16 Nov. 2022. 5. "Why Investing Pros Spooked by Market Manias Are Studying "Econophysics."" Fortune, fortune.com/longform/meme-stocks-stonks-gamestop-econophysics-market-maniasgiorgio-parisi/.
7. Bonabeau, Eric, and Christopher Meyer. "Swarm Intelligence: A Whole New Way to Think about Business." Harvard Business Review, 1 May 2001, hbr.org/2001/05/swarmintelligence-a-whole-new-way-to-think-about-business