USING ARTIFICIAL SWARM INTELLIGENCE TO IMPROVE PREDICTIONS OF REAL ESTATE LISTINGS

Through application of cutting-edge AI technology, Swarmalytics Labs has uncovered patterns in residential listings that can change the real estate market.

Summary

One of the greatest challenges in the real estate market is identifying who will likely list their home for sale in the near future. Listings are sporadic events often separated by several years of inactivity. Therefore, any method or algorithm that could reliably predict these events has the potential to generate immense value. For this reason, Swarmalytics Labs has recently concluded a series of test analyses of real estate activity. The purpose of these tests was to determine whether new analytic technologies could identify subtle predictive elements among the thousands of data elements related to any given dwelling unit.

The technology used to explore this massive volume of data is **artificial swarm intelligence**, an analog of the natural collective intelligence observed in flocks of birds, schools of fish, swarms of bees, etc. In artificial swarm intelligence, thousands of intelligent agents simultaneously search a data landscape for patterns and communicate their findings to each other. When applied to predictive analytics, the swarm of intelligent agents subdivides a problem into homogeneous segments, creates a solution for each segment, and finally reunites the solutions into an ultimate equation that is greater than the sum of its parts.

Introduction

In pursuit of real estate listings, many real estate agents rely on brute-force methods such as mass postcard mailings or thousands of phone calls and emails to determine which owners are in the market to sell their property. This is an extremely inefficient strategy, where the overwhelming majority of the marketing falls on unresponsive individuals. Targeting real estate listing solicitations holds the promise of far superior economics in the form of less unproductive effort and fewer materials wasted.

Background

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[1], and the 1992 publication of Marco Dorigo's *Optimization, Learning, and Natural Algorithms*, which laid the groundwork for ant colony optimization algorithms[2]. 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[3], which is also used to navigate their pilots to airport gates most efficiently[4]. 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*[5], 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[6].

In 2021, there were 7.2M real estate transactions in the United States with a record-setting $2.9T in total value[7]. The majority of these transactions are handled through real estate agents. It is common for agents to split their commission between the buyer's side and seller's side. Solicitation of listings is the major expenditure for the listing agent.

Swarmalytics

The volume of data being gathered daily continues to expand at a geometric rate. In 2020, over 64 zettabytes (or 64 billion terabytes) of data were captured worldwide, a figure projected to triple by 2025[8]. This includes massive amounts of data from the real estate world, leading to what has been called the "third phase of PropTech"[9]. Examples of available data for a typical single-family dwelling include descriptions of the property itself (lot size, construction date, number of bedrooms and bathrooms) but also neighborhood demographics, regional economic data, local crime statistics, agent and price histories, proximity to emergency services, local traffic flow data, etc. This can amount to tens of thousands of values associated with a single property. 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.

Model Development

Swarmalytics has massive amounts of data on every residence in the United States. This data has been analyzed via proprietary artificial swarm intelligence, resulting in a model to predict which individuals are most likely to list their homes based on behavioral and demographic characteristics and local market conditions. The model is the result of 7.4 million candidate equations generated at high speed over the course of 24 hours and then combined through the swarming process. The strength of this model is significant; the top 5% of scored individuals are 5X more likely to list their home in the next 90 days compared to the average among the population studied. Performance of the model was similar between the training and validation subsets, indicating that this model of predicting future home listings is highly reliable.

Swarmalytics

Performance of Swarmalytics' Listing Prediction Model
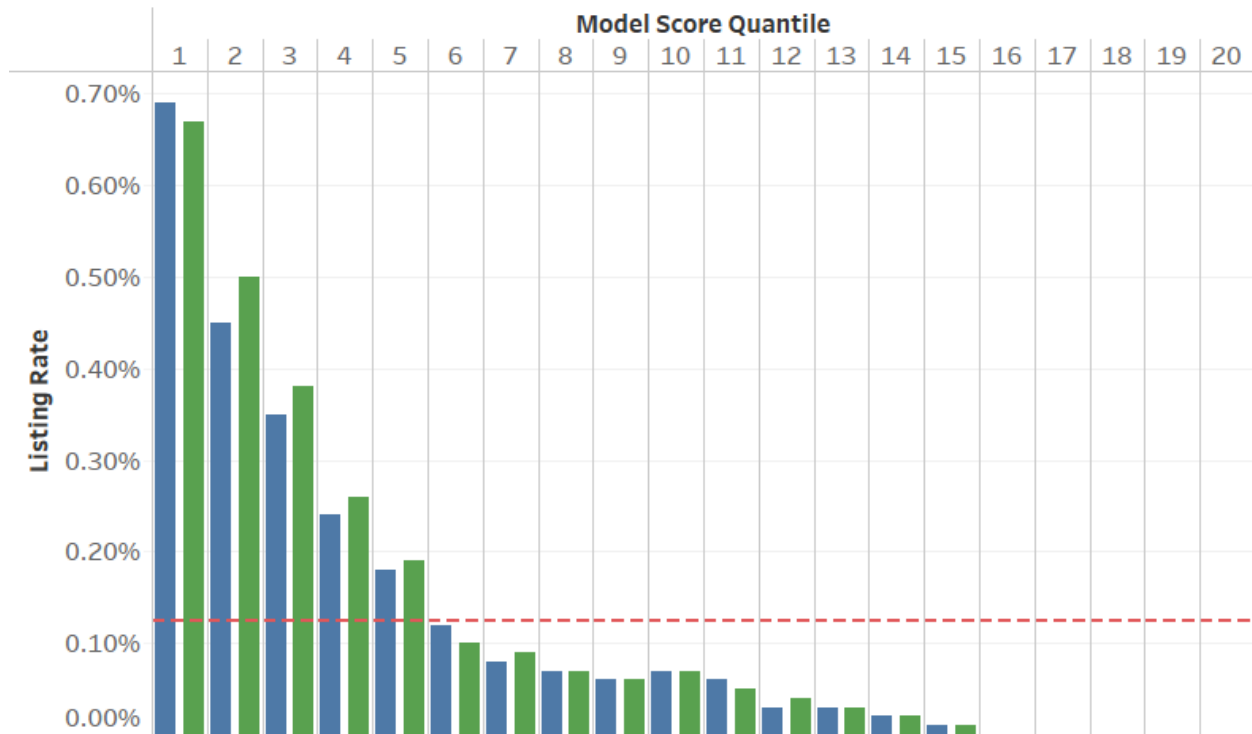
The overall listing rate of the data used to construct the listing model was 0.13% (shown by the dotted red line). When the data is scored by the model, the listing rate among the top 5% of scored data points (Model Score Quantile 1) is 0.67%, over five times higher than the overall rate. Note the similarity in performance between the training and validation subsets in each quantile, indicating that this is a very reliable model.

Case Study

To further demonstrate the effectiveness of this model, a subset of the Swarmalytics real estate database was created that included all real estate records for San Diego County, California. This area has a population of approximately 3.3 million residents and 528,276 single-family homes. It is worth noting that the average retail sale price for these homes is $1.1M, making each listing worth roughly $33,000 in listing commission. This subset of data served as the control group for testing. Separately, a dataset of 3,988 leads was created including four months of historical data. These leads were scored according to Swarmalytics' model. The test's experimental group was created after eliminating 538 leads that had already been listed and then creating a subset of the top 50% of scored properties for a total of 1,715 residences. The control and experimental groups were then monitored for 90 days. After this observation period, the final results were tallied for both groups. In the overall San Diego County housing market, 2,224 homes were listed for sale for a listing rate of 0.42%. However, 103 of the leads that scored in the top

50% according to Swarmalytics model were listed in the same three-month period. This demonstrates that **Swarmalytics' model was able to predict which single-family homes would be listed in the next 90 days with a success rate that is 14 times greater than random selection.**

Data Description	Test Group	Residence Count	Residences Listed in 90 Days	90-Day Listing Rate
San Diego County Single-Family Residences	Control	528,276	2,224	**0.4%**
Top 50% of Leads by Swarmalytics Model Score	Experimental	1,715	103	**6.0%**

Conclusion

Any organization able to predict which homes will be listed for sale stands to gain a substantial advantage over the competition. The proliferation of data covering every aspect of every home in the United States makes these predictions attainable, but only through the use of the most advanced analytic techniques, chief among them being artificial swarm intelligence. Swarmalytics is leading the way in this field and has already created a predictive model that is able to differentiate single-family homes that are likely to list in the next 90 days from those that are not.

About Swarmalytics

Swarmalytics, Inc. 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 in industries from real estate to e-commerce to healthcare. To learn more about Swarmalytics, please visit www.swarmalytics.com.

Swarmalytics

References/Citations

1. 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.
2. M. Dorigo, 1992. Optimization, Learning and Natural Algorithms, PhD thesis, Politecnico di Milano, Italy.
3. "What Can Ants Teach Us?" CBS News, www.cbsnews.com/news/what-can-ants-teach-us/. Accessed 16 Nov. 2022.
4. "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-manias-giorgio-parisi/.
6. 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/swarm-intelligence-a-whole-new-way-to-think-about-business.
7. Malone, Thomas. "2021: A Record-Breaking Year for Real Estate Transactions." CoreLogic®, 18 May 2022, www.corelogic.com/intelligence/2021-a-record-breaking-year-for-real-estate-transactions/. Accessed 16 Nov. 2022.
8. Statista. "Data Created Worldwide 2010-2025 | Statista." Statista, Statista, 7 June 2021, www.statista.com/statistics/871513/worldwide-data-created/.
9. Xu, Louisa. "Modernizing Real Estate: The Property Tech Opportunity." Forbes, www.forbes.com/sites/valleyvoices/2019/02/22/the-proptech-opportunity/?sh=ca03d7e58264. Accessed 16 Nov. 2022.

Swarmalytics