Machine Learning for Mortgage Lenders

Posted By **Radhika Ojha** On December 14, 2018 @ 4:09 pm In Daily Dose,Featured,News,Technology | Comments Disabled

In this Q&A, Michael Britvan, Managing Director Loan Sale and Asset Sale group at Mission Capital and Allison Israel, Product Manager of Mission Capital give insights into how machine learning and artificial intelligence will have a broad impact on lending operations.

How do you see artificial intelligence and machine learning impacting the mortgage space?

Israel: There are various applications for artificial intelligence across the mortgage industry, but one area where we're already seeing machine learning make an impact is the analysis of loan portfolios.

When banks explore the sale of loan portfolios in the secondary market, they produce data tapes containing relevant loan, collateral, and borrower information from their servicing systems. Field names in these tapes frequently vary by servicing platform as there is currently limited industry standardization. For example, a data field in one loan tape might refer to "Origination Date," while another shows "OrigDate" and a third is "Loan Origination Date." Although each of these fields refers to the same thing, the fact that they are labeled differently means that an analyst looking to load a model might spend considerable time deciphering column headers and normalizing data.

Machine learning has the power to take this manual process and perform it automatically. For example, it is able to recognize that "OrigDate" means "Origination Date." Additionally, when the system is processing a new tape and finds a term it doesn't recognize, it uses natural language processing to parse the word and find the closest match. The more tapes we put through the system, the smarter it gets. A few months ago–when we first deployed the system internally–it generally recognized around 40 percent of the fields. But, as it learns more, and processes a greater number of tapes, we expect that number to climb closer to 90 percent.

Would you say that the greatest benefit of machine learning is time savings?

Israel: While time savings is an important factor, having standardized field names from the machine learning model also allows us to apply a standard set of "rules" within the same software. For example, with all tapes using the term "Origination Date," we can tell the system that "Origination Date" must come before "Maturity Date," and it will flag any loans that don't comply with the rule. We currently have about 250 rules, and they are instrumental in enabling us to improve data integrity by catching data issues programmatically.

Conventionally, analysts have spent up to 80 percent of their time in Excel normalizing data, validating information in the tapes, and resolving errors. This results in very little time to analyze the value and potential of the portfolios at hand. With newly developed software, we're leveraging machine learning to flip the scale and enable analysts to spend less time manually manipulating data tapes and more time on the actual analysis.

Across the industry, loan analysis and trading are made infinitely more efficient by introducing machine learning models and enhancing those models with historical big data. The key to leveraging big data is the ability to normalize it first.

What are the other benefits mortgage professionals realize from this technology?

Britvan: The technology empowers all mortgage professionals to validate, analyze, and visualize data more efficiently. Depending on the user and firm, this can translate into a range of different benefits.

Banks leveraging this technology might be able to gain better insight into their portfolios. By cleaning up data and eliminating errors, they are also better able to manage their service providers. For example, with a better handle on their portfolio, it will become easier for banks to spot-check loan servicers to ensure accurate reporting and potentially even audit remittances.

Investors acquiring whole loans are able to spend more time on analysis and less time cracking tapes and stratifying portfolios.

Q: Do you think these innovations will have a broader impact on the whole loan sale market?

Britvan: Over the past decade or so, there's been a significant shift in the perception of trading whole loans on the secondary market.

Ten or twelve years ago, selling loans on the secondary market was often an indicator that the seller had a problem on their hands, and the decision to sell stemmed from a desire to remove the problem from their books. That perception has changed. Today, the speed of transactions has increased, while the number of participants in the secondary market for whole loans has climbed significantly. Whole loans are a relatively liquid asset, and many banks routinely tap the secondary market to manage their loan portfolio.

We expect technology to increase efficiency in analyzing loan portfolios which should, in turn, expand the universe of buyers in the secondary market. Right now, most buyers considering entering the market rely on an analyst to clean and validate data prior to loading a model. With the strides we're making in producing tools that clean up the data automatically, it allows investors to focus on finding value rather than allocating resources to data manipulation.

Do you think there are other notable tech trends that will have a significant effect on the secondary loan sale market?

Britvan: One area that has a lot of untapped potential is the incorporation of big data into mortgage analysis. When analyzing a loan portfolio, the quality of the valuation we can produce is often limited by the quality of data we receive. Key data points that are stale or absent require that assumptions be made.

By taking a big data approach to updating stale data or making assumptions, we can improve our estimates. For example, if we're analyzing a multifamily property, we could leverage things like demographic trends, market occupancy, existing and future inventory, and housing data to make assumptions regarding a property's current and projected future occupancy. This means that we are no longer just filling in missing or stale data but are also using historical trends to predict the market. This introduces brand-new inputs into our models that would have previously been unavailable without the breadth of data at our fingertips today.

Big data analysis of external factors, combined with proprietary market knowledge gleaned from our whole loan trading activity, will provide a better basis for secondary market participants to analyze loan tapes. We expect the industry to make significant strides in incorporating third-party data into their analysis in the years ahead.
