Machine Learning POC for a large Media company Case Study

SUMMARY

A large listed media company of India wanted to evaluate the applicability of predictive machine learning models to be able to predict the ratings for proposed TV programs across different time slots. They wanted to do a POC which would provide them with measurable data for building a business case.

THE CLIENT'S CHALLENGE

TV ratings are the basis for billions of dollars' worth of advertising transactions every year between marketers and media companies. Reliable forecast of ratings can help the media industry players make faster, more accurate and less subjective decisions. It also helps in measuring success of TV shows.

The customer being a large media house had to roll out programs within aggressive time schedules. The customer was making lot of investments to maintain a team of developers to write code manually for Data Verification, Analysis and model building. A high turnaround time running into months was required to complete the task of model building. There was also a dearth of people having niche skills of Machine Learning and Analytics for building ML models.

Customer wanted to make use of new automated technologies based on AI/ML to build predictive models which could analyze TV ratings from historical data and predict the future ratings for new programs to be launched. This would give them a competitive business edge and would directly enhance their revenues.

ACTION PLAN

Striatum was selected as the partner of choice for developing a proof of concept (POC) to build a ML model to predict future ratings across different time slots for a genre of TV program. A key mandate was to make use of an automated tool/platform which cuts down on the need for manual development effort for model building. The objective of the POC was to compare the predicted results with the actual ratings so as to measure the accuracy of the ML models.

TV rating and viewership data for 3 months for 3 different years were to be used for the POC.

Striatum built the POC within 2 weeks using one of our cutting-edge technology platforms – Sparkflows ®. The POC covered the following:

- Data pre-processing and analysis which included finding the distribution of program ratings across different timeslots, handling of outliers, feature engineering and finding the co-relation matrix.

- Data modelling and prediction – Random Forest and Gradient Boost Tree (GBT) models were used for building the predictive models

- Graphical Dashboards to visualise the trends observed and the actual vs predictions

Sparkflows ® enabled building the workflows for data pre-processing, analysis and predictive models using a user-friendly UI and built in nodes/processors and this helped us to complete the POC successfully within the time frames with good accuracy scores

RESULT

The accuracy score of the predictive models built as part of the POC was quite high. This gave confidence to the client that such models could be applied by them to predict future ratings of TV programs using automated AI/ML models with the help of powerful platforms like Sparkflows ®. As next steps the client started work on building an internal business case for making such of such tools and technologies.