Publication: IDA
Date: 3 January 2014
Headline: SMU team wins round 1 of DSM's Data Innovation Challenge

SMU team wins round 1 of DSM's Data Innovation Challenge

January 3, 2014

Their predictive sales forecasting model gets the judges' nod with best accuracy score.

How can a company accurately project its sales volume for a particular segment, in a particular region, for a particular year? This was the challenge facing the six teams taking part in Round 1 of the Data Innovation Challenge presented by the Dutch company DSM Engineering Plastics.

Launched in June 2013, the Data Innovation Challenge provides a platform for organisations to crowd-source and develop analytics solutions to address complex data-driven business challenges. It brings together user enterprises with data providers, data scientists, ICT companies who will collaborate with research institutes and institutions of higher learning to develop proof-of-concepts and test out the prototypes or working models to address the user enterprises' business challenges.

DSM Engineering Plastics, which supplies high-performance engineering thermoplastic solutions, is one of the user enterprises offering specific business challenges and datasets for the competition. It wanted to find a way to accurately project the 2013 sales volume for a certain segment of the business in the Asia Pacific, based on 2011 and 2012 results.

The winning solution came from Team Davidlowjw, which comprised Research Engineer Mr David Low and Research Scientist Dr Richard Oentaryo from the Living Analytics Research Centre, Singapore Management University.

The team members adopted a three-phase approach to the problem. First, they conducted an exploratory data analysis by visualising the trends in sales figures and looking at these for individual market segments and product types to discover correlations. Based on this, they devised a set of input variables (both qualitative and quantitative) that could help in the prediction of the sales figures. They then tried to "learn" the relationship between the input variables and the sales figures using a forecasting model based on a combination of machine-learning methods.

The team also created an internal validation dataset using the most recent three months' data, in order to evaluate and fine-tune the model for submission for the Data Innovation Challenge.

Team DavidLowjw went on to top this phase of the Data Innovation Challenge. According to Ms Clara Lee, Director, Business Information at DSM Engineering Plastics and a

member of the judging panel, it had the best accuracy score and also put in obvious effort to analyse the data, putting a lot of thought into what data to use, and what not to use, for their model. The team also used business-friendly language in their presentation and clearly explained how technology could be applied. "The potential of using the model is high," said Ms Lee.