A base de dados contava originalmente com 790 instâncias, porém apenas 500 foram disponibilizadas de forma pública devido a confidencialidades da página do Facebook em questão. Todas as instâncias contam com 19 atributos, após a criação de um modelo de previsão de performance da postagem, 7 desses atributos foram considerados como entrada e os outros 12 como medidas de performance.

Os 7 atributos de entrada são: a quantidade de likes totais da página no momento da postagem em questão, o tipo de postagem (foto, vídeo, link ou status), a categoria (atributo interno da empresa adicionado manualmente), mês, dia da semana e hora da postagem, cada um em um atributo separado e se o post foi pago para ter mais publicidade. Os outros 12 atributos de performance são: quantas pessoas o post atingiu, quantas vezes a postagem foi vista, quantos usuários interagiram com a postagem ao todo, quantidade de pessoas que interagiram com a postagem sem “compartilhar com outras pessoas” através de histórias, quando vezes o conteúdo da postagem foi clicado sem “compartilhar com outras pessoas” através de histórias, quantas vezes a postagem foi vista por pessoas que curtem a página em questão, quantas pessoas que curtem a página viram a postagem, quantas pessoas que curtem a página e interagiram com a postagem, quantidade de comentários, curtidas, compartilhamentos e o total desses últimos 3 atributos. Todos esses atributos, com exceção da categoria que foi adicionada manualmente por um funcionário da empresa, foram retirados das estatísticas da página da empresa de cosméticos diretamente do Facebook.

Após o processo de mineração utilizando SVM, foi gerado um modelo para cada um dos 12 atributos de performance utilizando os 7 atributos de entrada. Os resultados dos modelos foram comparados com os valores reais retirados dos dados através de diferença absoluta e percentual, além da média de erro absoluto percentual para cada um dos 12 atributos. Para facilitar a análise dos atributos de saída, estes foram divididos em visualização e interação, sendo que os de interação tiveram como melhor resultado 27% de erro e o melhor de visualização com 37% de erro. Especulasse que os erros maiores com relação a previsão de visualização se devem a maiores fatores fora do controle dos dados e dos modelos, como a aleatoriedade do algoritmo de feed do Facebook.