Loan Credit Risk Prediction

Fachrur Rozi, Rifki Adrian, Imam M. Rachbini

Maret 2018

# **Deskripsi Masalah**

Banyak sekali risiko yang dihadapi dalam dunia perbankan, salah satu skenario terburuk adalah kehilangan modal dikarenakan pemberian pinjaman/kredit yang tidak dapat dikembalikan oleh peminjam.

Maka dari itu, analisis resiko kredit adalah sebuah masalah krusial yang dihadapi oleh bank dan dapat membantu mengevaluasi apakah aplikasi pinjaman seorang pelanggan dapat diterima atau tidak. Penyelesaian isu ini diharapkan dapat membantu bank untuk meminimalisir risiko kehilangan modal, dan dapat meningkatkan volume kredit.

# **Dataset**

Dataset yang kami gunakan berasal dari Microsoft Azure dan bisa diakses pada tautan berikut <https://github.com/Microsoft/r-server-loan-credit-risk/blob/master/Data/> . Dataset ini memiliki kurang lebih 100.000 baris data. Pada dataset ini, terdapat banyak features yang bisa digunakan sebagai internal node pada Decision Tree nantinya. Sedangkan Feature yang akan dijadikan sebagai output leaf pada Decision Tree ini adalah loanStatus, yang mana domain dari output ini hanya ada 2, yaitu default dan current.

Default adalah output bagi mereka yang tidak diberikan pinjaman oleh lending institution, sedangkan current adalah output bagi mereka yang diberikan pinjaman oleh lending institution.

# **Algoritma Decision Tree CART (Classification And Regression Trees)**

Classification and Regression Trees adalah sebuah term dari Decision Tree algorithms yang dapat digunakan untuk masalah prediksi model dengan classification atau regression.

Reprsentasi dari CART adalah binary tree. Internal node dari tree adalah satu input variabel yang ingin diprediksi dan split point dari variabel tersebut. Leaf nodenya adalah output / label dari variabel tersebut.

Model Decision Tree yang akan kami buat untuk menyelesaikan masalah di atas adalah sebuah model yang input variabelnya (internal node) adalah fitur-fitur yang akan dipilih dari dataset, dan yang menjadi output variabel (leaf node) adalah nilai boolean apakah orang tersebut bisa mengajukan kredit bank atau tidak.