Analisis kinerja one dimensional naive bayes sebagai metode imputasi data masalah asuransi = Performance analysis of one dimensional naive bayes as a data imputation method for insurance problems

Natalia Aji Yuwanti, author

Abstrak

Metode machine learning sangat banyak digunakan dalam membantu pekerjaan manusia. Tidak semua data seperti yang diharapkan. Kebanyakan data memiliki missing value. Data yang memiliki missing value harus ditangani dulu pada tahap pra pengolahan, salah satunya adalah dengan cara imputasi missing value. Pada penelitian ini, dilakukan analisis kinerja One-Dimensional NaÃ¯ve Bayes sebagai metode imputasi data masalah asuransi mobil dan keselamatan berkendara. Berdasarkan hasil simulasi menggunakan SVM didapatkan hasil yang sama untuk imputasi menggunakan modus dan One-Dimensional NaÃ¯ve Bayes pada data Car Insurance yaitu 1,00. Setelah itu dilakukan telaah lebih lanjut ternyata imputasi setiap missing value dengan modus dan prediksi imputasi dengan One-Dimensional NaÃ¯ve Bayes persis sama. Pada data Safe Driver, imputasi dengan modus menghasilkan akurasi 0,86 sedangkan imputasi dengan One-Dimensional NaÃ¯ve Bayes menghasilkan akurasi 0,85. Hasil ini menunjukkan bahwa metode imputasi missing value dengan modus masih sangat direkomendasikan untuk tahap pra pengolahan data pada machine learning.

.....Machine learning methods are very widely used in helping human work. Not all data is as expected. Most data have missing values. Data which has a missing value must be handled first at the pre-processing stage, one of which is by imputation of the missing value. In this study, a One-Dimensional NaÃ¯ve Bayes performance analysis was performed as a data imputation method for car insurance and safe driver problems. Based on simulation results by using SVM obtained the same results for imputation using mode and One-Dimensional NaA ve Bayes on Car Insurance data that is 1,00. After that, a further study is carried out, apparently the imputation of each missing value by mode and the prediction of imputation with One-Dimensional NaAve Bayes are the same. In Safe Driver data, imputation with mode produces 0.86 accuracy while imputation with One-Dimensional NaAve Bayes produces accuracy of 0.85. These results indicate that the method of missing value imputation with mode is still highly recommended for the pre-processing data stage in machine learning.