

Abstrak

Perkembangan teknologi dalam pengumpulan dan penyimpanan data, maka muncul sebuah kebutuhan akan informasi penting dari data yang. Klasifikasi merupakan sebuah teknik yang dapat dilakukan untuk mengekstrak informasi penting dari data yang ada. *Rotation Forest* merupakan teknik klasifikasi metode *ensemble* yang menggunakan prinsip *feature extraction* menggunakan *Principal Component Analysis* (PCA) dan *decision tree* dalam membangun *base classifier*-nya. *Decision tree* dipilih karena *decision tree* sangat sensitif terhadap rotasi yang dilakukan pada data dengan membangun *tree* yang lebih sedikit. Hasil pengujian menunjukkan bahwa hasil performansi secara keseluruhan dipengaruhi oleh nilai jumlah *base classifier* (L) dan jumlah *subset* (K). Nilai K yang lebih kecil memberikan performansi yang lebih baik karena jumlah atribut yang besar dengan pembagi K yang lebih kecil akan mengoptimalkan kinerja PCA untuk data berdimensi besar. Sementara, nilai L yang lebih kecil mampu meningkatkan performansi karena jumlah *decision tree* yang dibangun sebagai *base classifier* mampu memberikan nilai akurasi yang lebih baik jika *tree* yang dihasilkan lebih sedikit. Dari hasil pengujian dapat dilihat bahwa, algoritma *Rotation Forest* memberikan hasil performansi yang lebih baik dalam memprediksi mahasiswa IT Telkom berpotensi *Drop Out* (DO) untuk data berdimensi besar dibandingkan metode *ensemble* lainnya seperti *LogitBoost* dan *Random Forest*.

Kata Kunci: *Rotation Forest, base classifier, decision tree, metode ensemble, feature extraction, Principal Component Analysis (PCA)*