# Random Forest

Logo, company name

Description automatically generatedRandom forest merupakan kumpulan atau kombinasi set yang menjadi satu sehingga membentuk forest (hutan atau kumpulan pohon) secara acak. Dari data yang acak ini diperoleh dengan cara yang sama seperti operator acak forest yang menghasilkan pohon. Algoritma ini nantinya dapat digunakan sebagai alat untuk melakukan pengklasifikasian dan regresi. Dari jumlah pohon yang ada nantinya akan dibentuk model hutan tertentu dari model pohon secara acak. Hutan tidak mengubah training set, tetapi menghasilkan kompetitif data [1].

Pengklasifikasian menggunakan metode ini memiliki fokus untuk meminimalisir kesalahan saat melakukan pengklasifkasian, sehingga dapat ditulis dengan 3 persamaan k ∈ (1, ..., K) sebagai kategori peristiwa, proporsi k yang ditunjukkan adalah *PkL* pada simpul kiri, dan *PkR* pada simpul kanan ditunjukkan proporsi k.

# Accuracy & Kappa Statistic

A picture containing text

Description automatically generatedAccuracy merupakan klasifikasi ketepatan record setelah dilakukannya pengujian hasil dari pengklasifikasian dengan hasil yang benar [2]. Fungsi dari accuracy ini sendiri yaitu untuk mencari nilai real dari suatu pengukuran secara absolut dan komparatif.pengukaran dengan rincian sebagai berikut:

Kappa Statistic merupakan analistik secara statistikal yang didasari dari kecocokan antar datanya, bisa dari penafsiran maupun derajat kesepakatan (*degree of agreement*) . analisis ditunjukkan bahwa kappa statistic memiliki kelas-kelas yang berbeda [2]. Kappa dianggap memiliki ukuran yang lebih akurat dibandinkan eprhitungan prediksi yang terjadi secara menyeluruh. Perhitungan ditulis sebagai berikut:

A picture containing diagram

Description automatically generated

# Cross Validation

Cross validation merupakan metode statistik yang fungsinya sebagai pengevaluasi dari kinerja suatu model atau algoritma yang kemudian dipisahkan menjadi 2 subset yaitu :

1. Data proses pembelajaran.
2. Evaluasi.

Timeline

Description automatically generatedCross validation menghasilkan akurasi dari proses training dan testing dari 2 data yang telah diklasifikasikan [3]. Pada penelitian kali menggunakan metode *10*-*fold cross validation* memiliki 10 lipatan (fold) dengan ukuran yang sama dengan hasil 10 subset data yang nantinya akan dilakukan evaluasi secara kinerja model.