**Kesimpulan Proyek**

Proyek ini bertujuan untuk melakukan analisis data cuaca di Seattle dan membangun model klasifikasi menggunakan algoritma Machine Learning untuk memprediksi kondisi cuaca berdasarkan atribut tertentu. Berikut adalah beberapa poin penting yang dapat diambil sebagai kesimpulan dari proyek ini:

1. Data Preprocessing: Proses awal proyek mencakup pembacaan dataset cuaca Seattle dari file CSV. Data preprocessing dilakukan dengan menghapus nilai kosong, menghapus kolom yang tidak diperlukan (kolom 'date'), dan menghapus data yang memiliki nilai duplikat.
2. Normalisasi Data: Atribut numerik dalam dataset, seperti suhu maksimum, suhu minimum, kecepatan angin, dan curah hujan, dinormalisasi menggunakan skala Min-Max.
3. Visualisasi Data: Data cuaca divisualisasikan untuk memahami relasi antara beberapa atribut numerik. Scatter plot digunakan untuk menampilkan relasi antara suhu maksimum dan suhu minimum, serta antara curah hujan dan kecepatan angin. Selain itu, distribusi target-class (jenis cuaca) ditampilkan dalam diagram pie.
4. Pembagian Data: Dataset dibagi menjadi train-set (70%) dan test-set (30%) menggunakan library **train\_test\_split**.
5. Model Machine Learning: Dalam proyek ini, digunakan CatBoost Classifier, yang merupakan salah satu algoritma klasifikasi yang kuat. Model CatBoost diubah dan dilatih dengan data train.
6. Evaluasi Model: Model yang telah dilatih digunakan untuk melakukan prediksi pada data test. Akurasi dari model ini diukur menggunakan metrik **accuracy\_score** dan ditampilkan dalam bentuk persentase.
7. Kesimpulan: Model CatBoost Classifier memberikan akurasi yang baik dalam memprediksi jenis cuaca berdasarkan atribut numerik. Akurasi ini dapat digunakan untuk memberikan informasi prakiraan cuaca yang lebih baik.

Proyek ini memberikan contoh aplikasi Machine Learning dalam memprediksi jenis cuaca berdasarkan data historis. Hasilnya dapat digunakan untuk mendukung keputusan dan perencanaan yang berkaitan dengan cuaca di wilayah Seattle.