|  |  |  |
| --- | --- | --- |
| **NO** | **POINT** | **KETERANGAN** |
| 1 | Judul | Studi Komparatif Algoritma Clustering dan Klasifikasi Pada Model Prediksi Tingkat Kehadiran Pegawai Negeri Sipil |
| 2 | Sumber Data | 1. Wawancana
2. Kuisioner
3. Data Base Kehadiran
 |
| 3 | Algoritma Data Mining | *Clustering*1. K-Means
2. DBScan
3. K-Medoids

Klasifikasi1. SVM (Support Vector Machine)
2. Decision Tree
3. Naïve Bayes

Dilakukan juga *hyperparameter tuning*. |
| 4 | Data | 1. Tabel
* NIP
* Nama Pegawai
* Domisili
* Usia
* Jabatan
* Manajemen Waktu
* Lingkungan Kerja
1. Data kehadiran PNS DKI Jakarta Periode 01 Januari – 31 Desember 2022
 |
| 5 | Tujuan | 1. Membandingkan hasil analisis kehadiran dengan tiga metode atau jenis Algoritma. Yaitu, K-Means, DBScan, dan K-Medoids.
2. Mendapatkan hasil analisis faktor yang paling mempengaruhi tingkat kehadiran PNS.
3. Dari hasil pengelompokkan data yang didapatkan oleh algoritma clustering terbaik, maka akan dilakukan klasifikasi dengan membandingkan 3 metode yaitu SVM, Decision Tree, dan Naïve Bayes.
4. Hasil dari performa algoritma klasifikasi terbaik akan digunakan untuk memprediksi label baru pada data inputan baru yang belum pernah dilihat oleh model algoritma sebelumnya.
5. Hasil dari pengelompokan dan prediksi tersebut dapat digunakan sebagai rujukan oleh pimpinan dalam mengambil keputusan dalam hal ini penempatan pegawai maupun perbaikan hal-hal terkait.
 |
| 6 | Alat Kerja | Python, Google Colab |
| 7 | Jurnal | Minimal 3-5 rujukan utama1. Komparasi Clustering (2)
2. Komparasi Klasifikasi (2)
3. Combine the Clustering & Classification method

**Gap Analisis:**  |