Nama : Atthariq Haykal Putera

NPM : 15210005

Prodi : Teknik Informatika

Matkul : Data Mining

**I**

**Jawaban:**

**1. B. Class**

**2. F. Cluster**

**3. I. Partisi**

**4. C. Dendogram**

**5. G. Maximum Distance**

**6. A. DBSCAN**

**7. J. AHC**

**8. H. Unsur**

**9. D. Confidence**

**10. E. Itemset**

**II**

**Jawaban**

1. **CRISP-DM (Cross Industry Standard Process for Data Mining).**
2. **Clustering, Regretion, Association Rule**
3. **Hierarchical dan Non Hierarchical**
4. **Euclidean Distance**
5. **Padat, bisa dibilang memiliki kepadatan yang bervariatif**

**III**

1. **Association Rule, merupakan sebuah metode dalam data mining dimana proses tersebut dilakukan untuk mencari hubungan atau mengidentifikasi hubungan suatu variabel dengan variabel yang lain dalam dataset**
2. **Terdapat 2 Parameter**

* **Support**

**Menguji seberapa sering suatu itemset bisa muncul dalam dataset**

* **Confidence**

**Menguji berapa kemungkinan itemset B bisa muncul dalam proses yang melibatkan itemset A**

1. **Manfaat**

* Mempermudah startegi bisnis
* Mempercepat analisis

1. **Penjelasan**

Dari rules diatas diambil kesimpulan 'kemungkinan itemset B (Diapers) muncul dalam transaksi itemset A ( Susu Formula, Tisu Basah) mencapai 80% dan bisa dikategorikan tinggi dan Untuk keseluruhan transaksi kemungkinan muncul 3 item diatas (Susu Formula, Tisu Basah, Diapers) mencapai 60% dari keseluruhan dataset transaksi yang terjadi

1. **Strategi**

Memberikan penawaran gratis diapers apabila membeli susu formula dan juga tisu basah secara bersamaan dalam 1 kali transaksi

1. Tahapan Evaluation dan Deployment

**Evaluation**

Untuk tahapan ini dilakukan perhitungan manual dengan menggunakan algoritma yang sudah ditentukan sebelumnya, biasanya bisa menggunakan excel dan data yang digunakan hanya sampel saja dari dataset yang ada

**Deployment**

Tahapan ini mengacu pada penerapan pyhon untuk data mining, disini juga sama menerapkan algoritma yang sudah ditetapkan kemudian diterapkan di python biasanya menggunakan Google Colab atau juga bisa menggunakan python di visual studio code

1. K-Means Clustering

Merupakan sebuah algoritma yang digunakan dalam data mining dimana, dalam proses clusterisasi akan di klusterisasi berdasarkan seberapa banyak kluster yang ditentukan sebelumnya dan ini termasuk ke dalam Unsupervised Learning