PENGOLAHAN CREDIT CARD CUSTOMER DATA DENGAN PEMODELAN UNSUPERVISED LEARNING MENGGUNAKAN METODE K – MEANS DAN AGGLOMERATIVE HIERARCHICAL

**Wilando Putra Y.1, Aldian Farhan P.2.**

1Teknik Informatika, Institut Teknologi Sumatera,Indonesia

2Teknik Informatika, Institut Teknologi Sumatera,Indonesia

Corresponding author: Aldian.119140080@student..itera.ac.id

**ABSTRACT** Teknologi pada zaman sekarang telah sangat berkembang, banyak fitur yang diciptakan melalui teknologi yang digunakan untuk memenuhi kebutuhan. *Unsupervised Learning* adalah salah satu tipe algoritma machine learning yang digunakan untuk menarik kesimpulan dari dataset. Metode ini hanya akan mempelajari suatu data berdasarkan kedekatannya saja atau yang biasa disebut dengan clustering. Metode unsupervised learning yang paling umum adalah analisis cluster, yang digunakan pada Analisa data untuk mencari pola – pola tersembunyi atau pengelompokan data. Pada jurnal ini metode algoritma digunakan untuk pengolahan data kartu kredit pelanggan.

**INDEX TERMS** Unsupervised Learning, Machine Learning, K-Means Cluster, Dataset.

# I. INTRODUCTION

Seiring dengan berkembang nya zaman, teknologi yang dibuat juga semakin canggih dan lebih banyak digunakan oleh manusia. Pada zaman sekarang contoh penggunaan kemajuan teknologi yaitu untuk memenuhi kebutuhan hidup per orang, semakin maju teknologi maka manusia lebih memilih untuk memenuhi kebutuhan hidup nya dengan cara yang cepat, mudah, dan praktis. Apalagi dimasa pandemi seperti ini masyarakat lebih memilih untuk melakukan kegiatan atau mencari kesibukan hanya melalui tempat tinggal masing – masing, begitu juga untuk memenuhi kebutuhan bisa dengan cara menggunakan smartphone untuk melakukan pembelanjaan. Kartu kredit juga merupakan pemanfaatan dari kemajuan teknologi, kartu kredit digunakan sebagai menjadi alat pembayaran yang sangat popular dikalangan perkerja atau orang yang sudah memiliki penghasilan rutin. Hal ini terbukti dengan adanya data dari Bank Indonesia bahwasanya jumlah kartu kredit yang beredar per januari 2014 mencapai 15,12 juta ([www.detik.com](http://www.detik.com)).

# II. STUDI LITERATUR

1. Unsupervised Learning

*Machine Learning* menggunakan Teknik untuk menangani data besar (big data) dengan cara yang cerdas untuk memberikan hasil yang tepat. Berdasarkan Teknik pembelajarannya, tipe – tipe *machine learning* dapat dibedakan menjadi *supervised learning* , *unsupervised learning, semi supervised learning dan reinforcement learning. Supervised learning* merupakan salah satu Teknik *machine learning* yang menggunakan dataset (data training) yang sudah berlabel untuk melakukan pembelajaran pada mesin, sehingga mesin mampu mengidentifikasi label input dengan menggunakan fitur yang dimiliki untuk selanjutnya melakukan prediksi mauppun klasifikasi, sedangkan *unsupervised learning* adalah Teknik dengan menarik kesimpulan berdasarkan dataset yang merupakan input data *labeled response* (Z. A. Fikria, 2017).

1. K-means

Cluster Analysis merupakan salah satu metode objek mining yang bersifat tanpa latihan (unsupervised analysis), sedangkan K-Means Cluster Analysis merupakan salah satu metode cluster analysis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokan kedalam cluster yang lain. Tujuan pengelompokan adalah untuk meminimalkan objective function yang di set dalam proses clustering, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu cluster dan memaksimalkan variasi antar cluster. (Ediyanto)

1. Agglomerative Hierarchical

Pengelompokan hierarkis dan aglomerat adalah teknik penting dan mapan dalam pembelajaran mesin tanpa pengawasan. Skema pengelompokan aglomeratif dimulai dari partisi

kumpulan data menjadi node tunggal dan menggabungkan langkah demi langkah pasangan saat ini yang paling dekat

node menjadi node baru sampai ada satu node terakhir yang tersisa, yang terdiri dari seluruh kumpulan data.

Berbagai skema pengelompokan berbagi prosedur ini sebagai definisi umum, tetapi berbeda dalam cara

di mana ukuran perbedaan antar-cluster diperbarui setelah setiap langkah. (Müllner)

# III. METODE

Pada tugas besar ini, metode yang digunakan yaitu metode K-Means yang termasuk dalam pemodelan *Machine Learning* bertipe *Unsupervised Learning.* Algoritma K-Means erguna untuk clustering/mengelompokkan data kartu kredit kostumer berdasarkan histrori kostumer itu sendiri.

# IV. DATA SET

Dataset yang digunakan pada laporan tugas besar ini terdapat pada spesifikasi tugas. Kami menggunakan dataset mengenai Credit Card Costumer. Terdapat 6 atribut, yaitu Costumer Key, Rata-rata Batas kartu kredit, Total kartu kredit kostumer, Total bank yang dikunjungi kostumer,Total visit online dan Total calls made.

Table

Description automatically generated with medium confidence

# V. HASIL EKSPERIMEN

1. Avg Credit Limit

Chart, bar chart

Description automatically generated

Gambar diatas merupakan data jumlah maksimal kartu kredit yang dimiliki oleh seorang costumer bank tersebut, data ini diperlukan bagi sebuah bank untuk memprediksi kebutuhan costumer.

2. Avg Visits Online

Chart, bar chart

Description automatically generated

Gambar diatas merupakan data kostumer yang melakukan kunjungan secara online, data ini berguna bagi sebuah bank untuk mengetahui kelengkapan fasilitas yang dimiliki oleh bank tersebut dan apakah berguna bagi costumer atau tidak.

3. Total Calls Made

Chart

Description automatically generated

Gambar diatas merupakan data dari maksimal total panggilan yang dilakukan oleh costumer per hari nya. Data ini berguna bagi sebuah bank untuk mengetahui tingkat pelayanan dari bank tersebut.

# VI. PEMBAHASAN

Pada tugas besar ini menggunakan PANDAS PYTHON untuk membaca dataset yang sudah ada di dalam Credit Card Costumer Data.csv. Dibawah ini dilampirkan beberapa perintah yang digunakan pada PANDAS PYTHON.

Graphical user interface, text, application

Description automatically generated

Perintah diatas digunakan untuk mengimpor data yang ada pada dataset yang sudah didownload dari spesifikasi tugas besar,yang kemudian akan diolah.





Kedua perintah diatas digunakan untuk menghitung jumlah table horizontal dan vertical pada dataset.



Perintah diatas digunakan untuk mendesripsikan hitungan jumlah table, nilai tengah, nilai terendah sampai nilai tertinggi.

Berikut ini dilampirkan hasil penghitungan rata-rata data mengenai kostumer berdasarkan grafik dan plot dot.

Diagram

Description automatically generated

Dan juga dilampirkan hasil penghitungan rata-rata data mengenai kostumer berdasarkan warna.

Chart

Description automatically generated

Berikut gambar analisis cluster dengan boxplot

Calendar

Description automatically generated with medium confidence

Chart

Description automatically generated with low confidence

Gambar diatas merupakan perintah untuk menampilkan Agglomerative Hierarchical Clustering Diagram.

Graphical user interface, text, application

Description automatically generated

Gambar diatas merupakan perintah K-Means yang digunakan untuk clustering.

Hasil akhir yang diperoleh dari dataset yang diolah, data tersebut dapat digunakan sebuah bank untuk lebih melakukan pendekatan dengan costumer nya dengan cara mengembangkan solusi dari data terendah tiap rata-rata yang berhubungan dengan costumer.

# VII KESIMPULAN

Kesimpulan yang dapat diperoleh dari percobaan yang dilakukan yaitu :

1. Metode K-Means sangat cocok dan berguna untuk clustering dataset.
2. Metode Diagram Agglomerative Hierarchical untuk clustering dataset.
3. Pembuatan pengelompokan data sangat berguna bagi yang memiliki data tersebut guna meningkatkan kualitas, pelayanan, dan sebagainya.
4. Teknik unsupervised learning memiliki metode yang berguna untuk melakukan clustering data.

# REFERENCES

[1] Nurhayati, Busman, and R. P. Iswara, “Pengembangan Algoritma Unsupervised Learning Technique Pada Big Data Analysis di Media Sosial sebagai media promosi Online Bagi Masyarakat,” J. Tek. Inform., vol. 12, no. 1, pp. 79-96, 2019.

[2] Ediyanto, Muhlasah Novitasari Mara, Neva Satyahadewi, “PENGKLASIFIKASIAN KARAKTERISTIK DENGAN METODE K-MEANS CLUSTER ANALYSIS”, FMIPA UNTAN, Volume 02 , No. 2 (2013), hal 133 – 136

[3] Daniel Müllner, “Modern hierarchical, agglomerative clustering algorithms”.

.

.