

ETL PIPELINE DI PT SIRKADIAN INDONESIA SEHAT

I.M.K.D. Jaya¹, I.G.A.G.A. Kadnyanan², dan L.G. Astuti³

ABSTRAK

ETL (Extract, Transform, Load) pipeline adalah salah satu teknik pengolahan data yang penting dalam dunia bisnis dan teknologi informasi. ETL pipeline digunakan untuk memindahkan data dari sumber yang berbeda, melakukan transformasi data, dan memuat data ke database atau data warehouse tujuan. Dengan menggunakan ETL pipeline, data dapat diolah secara efisien dan efektif, sehingga informasi yang dihasilkan lebih akurat dan tepat waktu. Pada PT Sirkadian Indonesia Sehat, ETL pipeline ini digunakan untuk memudahkan data scientist mengambil data yang sudah diubah agar dapat membuat sistem rekomendasi dengan lebih baik lagi.

Kata kunci : Extract, Transform, Load, Pipeline, Sirkadian

ABSTRACT

An ETL (Extract, Transform, Load) pipeline is one of the crucial data processing techniques in the world of business and information technology. ETL pipelines are used to move data from different sources, perform data transformations, and load data into a target database or data warehouse. By utilizing ETL pipelines, data can be processed efficiently and effectively, resulting in more accurate and timely information. At PT Sirkadian Indonesia Sehat, this ETL pipeline is used to facilitate data scientists in retrieving transformed data to improve the recommendation system even further.

Keywords: Extract, Transform, Load, Pipeline, Sirkadian

1. PENDAHULUAN

Perkembangan teknologi yang semakin cepat di era globalisasi saat ini sudah tidak dapat dihindari lagi (Agustian, Salsabila, 2021). Hampir seluruh bidang sudah memakai teknologi yang menyebabkan metode konvensional secara perlahan mulai ditinggalkan, terutama pada teknologi basis data (Effendi, Wahidy, 2019) (Berliantara et al., 2017). Berbagai perusahaan berlomba-lomba untuk mengikuti teknologi yang ada dikarenakan peranan teknologi dapat membantu proses usaha (Nofitri, Irawati, 2019). Salah satu contoh perusahaan yang memanfaatkan teknologi yaitu PT Sirkadian Indonesia Sehat.

¹ Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana, Bukit Jimbaran, Badung, 80361, Badung-Indonesia, dwipajaya19@gmail.com

² Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana, Bukit Jimbaran, Badung, 80361, Badung-Indonesia, gungde@unud.ac.id

³ Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana, Bukit Jimbaran, Badung, 80361, Badung-Indonesia, lg.astuti@unud.ac.id

Submitted: 7 Oktober 2023

Revised: 26 Oktober 2023

Accepted: 26 Oktober 2023

Sirkadian merupakan sebuah startup yang bergerak di bidang kesehatan untuk memudahkan penggunanya dalam menerapkan pola hidup sehat. Saat ini, PT Sirkadian sedang membangun sistem rekomendasi makanan untuk para penggunanya menggunakan metode unsupervised learning. Unsupervised learning adalah teknik yang digunakan untuk mengambil kesimpulan berdasarkan dataset yang merupakan input data labeled response (Retnoningsih et al., 2020). Penggunaan metode unsupervised learning ini biasanya digunakan untuk mengelompokkan atau mengkategorikan data (Abijono et al., 2021). Dalam pembuatan sistem rekomendasi ini, diperlukannya data untuk melatih sistem yang akan dibuat. Adanya ETL Pipeline yang berfungsi untuk membentuk kembali data yang relevan dari berbagai database dan menjadi informasi yang berguna serta disimpan di data warehouse (Oslan, Kristanto, 2019). Data yang disimpan pada data warehouse sudah siap pakai untuk melatih sistem.

ETL (Extract, Transform, Load) adalah proses mengumpulkan, mengolah, dan menggabungkan data dari berbagai sumber dan disimpan kedalam data warehouse (Putra et al., 2019). Pada ETL ini juga terdapat proses pembersihan data seperti menangani redundansi, integritas, dan inkonsistensi data (Wijaya et al., 2016). Dengan menggunakan ETL pipeline, data dapat diolah secara efisien dan efektif dikarenakan dikumpulkannya data dari sumber yang berbeda dan membuatnya menjadi satu kesatuan data yang terintegrasi (Andriansyah, 2022). Saat ini, PT Sirkadian Indonesia Sehat sedang mengembangkan sistem rekomendasi pada aplikasinya. Pada aplikasi yang dibuat oleh PT Sirkadian Indonesia Sehat ini masih menyimpan data dari berbagai database. Hal ini akan memakan waktu dan proses yang cukup lama jika data diambil secara terpisah. Oleh karena itu, dibangun ETL Pipeline oleh penulis untuk mempercepat dan memudahkan dalam pembuatan sistem rekomendasi tersebut. Diharapkan dengan adanya ETL Pipeline ini, Data Scientist di PT Sirkadian Indonesia Sehat dapat dengan mudah mengambil data yang diperlukan dari berbagai database dan dapat menggunakannya untuk melatih sistem rekomendasi yang akan dibuat.

2. METODE PELAKSANAAN

2.1. Pelaksanaan Kegiatan

Pengembangan sistem ETL Pipeline yang dilakukan di PT Sirkadian Indonesia Sehat dilakukan menggunakan metode pengembangan agile development yang memungkinkan pengembangan yang cepat dan juga cepat untuk beradaptasi dengan perubahan.

2.2. Kebutuhan Fungsional Sistem

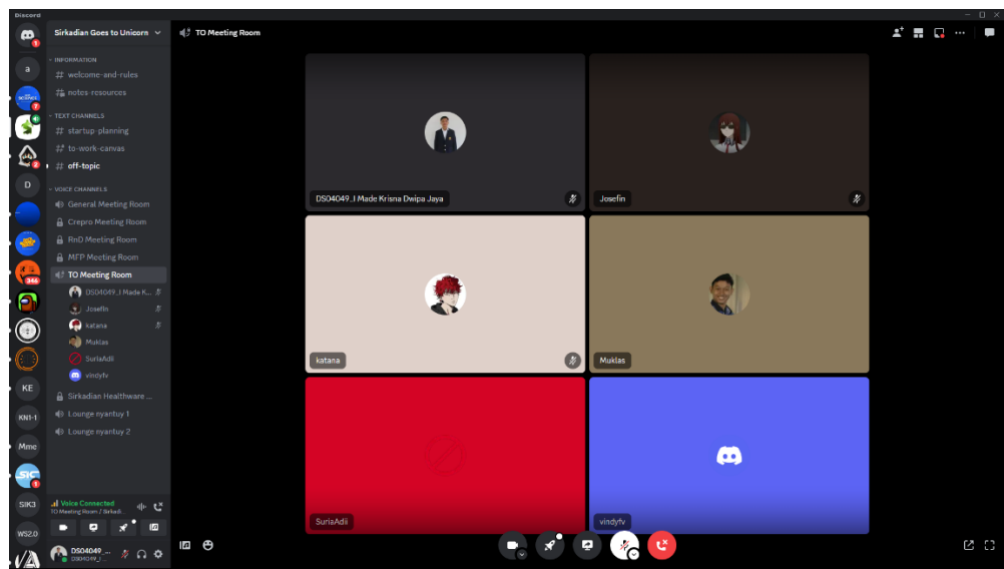
Pengembangan sistem ETL Pipeline untuk PT Sirkadian Indonesia Sehat memiliki beberapa fitur yang harus diselesaikan untuk memastikan sistem berjalan sesuai tujuan. Adapun beberapa kebutuhan fungsional yang dapat dilihat pada tabel di bawah.

Tabel 2.1. Daftar Kebutuhan Fungsional

No.	Fungsional
1	Mengekstrak data
2	Mentransform data
3	Memuat data

3. HASIL DAN PEMBAHASAN

Kegiatan pengabdian di PT Sirkadian Indonesia Sehat dilaksanakan pada tanggal 10 April sampai 08 Juni 2023. Hasil dari pengabdian selama dua bulan di PT Sirkadian Indonesia Sehat ini berupa sistem ETL Pipeline yang dibuat untuk memudahkan dalam pembuatan sistem rekomendasi.



Gambar 3.1 Proses Analisa Kebutuhan

```
{
  "ke": "Neva.Gerlach54@gmail.com",
  "user_id": "725a46cd-e051-4c12-bcc5-32cfd0a9195a",
  "hello": {
    "reply": "hello N
eva.Gerlach54@gmail.com"
  },
  "datetime": "2023-06-04 10:57:13"
}
{"ke": "Neva.Gerlach54@gmail.com",
"user_id": "725a46cd-e051-4c12-bcc5-32cfd0a9195a",
"hello": {
  "reply": "hello N
eva.Gerlach54@gmail.com"
},
"datetime": "2023-06-04 10:57:13"
}
```

Gambar 3.2 Tampilan ketika sistem dijalankan

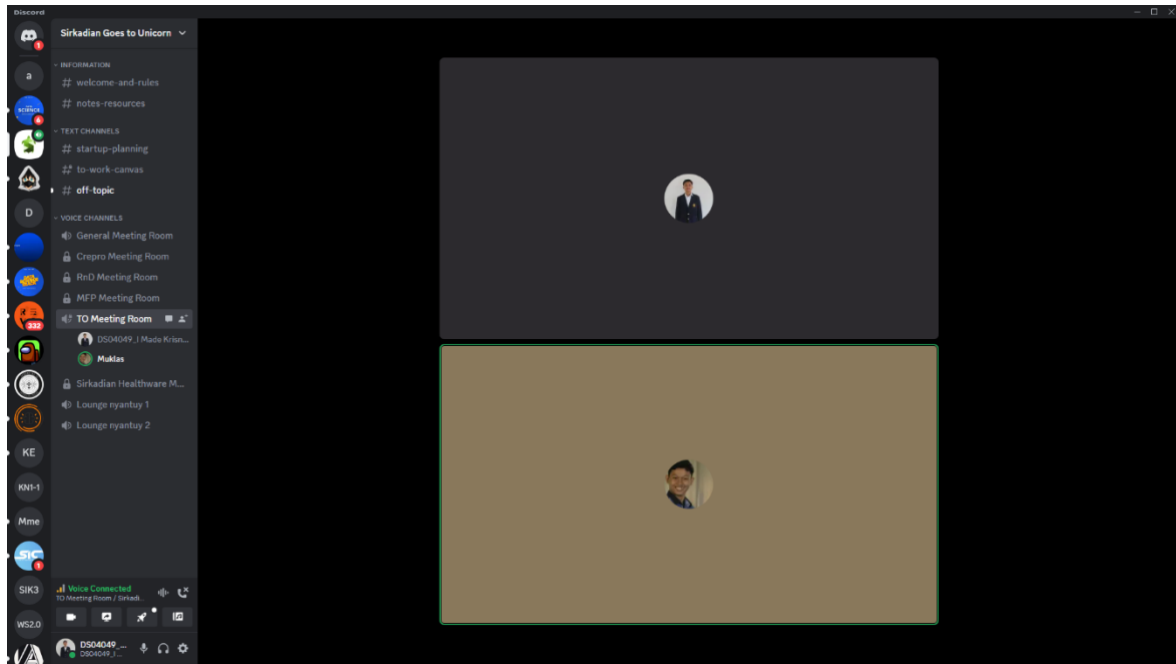
data_testing3 QUERY SHARE

SCHEMA DETAILS PREVIEW LINEAGE DATA PI

Filter Enter property name or value

☐	Field name	Type	Mode	Key	Collation
☐	ke	STRING	NULLABLE		
☐	user_id	STRING	NULLABLE		
☐	hello	RECORD	REPEATED		
☐	datetime	DATETIME	NULLABLE		

Gambar 3.3 Data berhasil tersimpan di data warehouse



Gambar 3.4 Sosialisasi ETL Pipeline dengan Mentor sekaligus Data Scientist PT Sirkadian Indonesia Sehat

4. KESIMPULAN

ETL Pipeline sudah berhasil dibuat dan diimplementasikan di PT Sirkadian Indonesia Sehat. Adanya ETL Pipeline dapat membantu data scientist di PT Sirkadian Indonesia Sehat dalam mengambil data yang sudah dikumpulkan pada satu data warehouse dan memudahkan pekerjaan mereka dalam membangun sistem rekomendasi.

UCAPAN TERIMA KASIH

Ucapan terimakasih disampaikan kepada PT Sirkadian Indonesia Sehat, atas diberikannya kesempatan terkait program pengabdian yang telah dilaksanakan, kemudian untuk Rektor Universitas Udayana, Ketua LPPM beserta staff. Ucapan terimakasih juga disampaikan kepada Ketua Program Studi Informatika, Bapak Dr. Ir. I Ketut Gede Suhartana, S.Kom., M.Kom., IPM., ASEAN.Eng., Ketua Pelaksana PKL Bapak I Wayan Supriana, S.Si., M.Cs., Dosen Pembimbing Bapak I Gusti Agung Gede Arya Kadyanan, S.Kom, M.Kom dan Ibu Dr. Dra. Luh Gede Astuti, M.Kom. yang telah membantu serta memberikan arahan dalam pelaksanaan program ini.

DAFTAR PUSTAKA

- Abijono, H., Santoso, P., and Anggreini, N. L. (2021). ALGORITMA SUPERVISED LEARNING DAN UNSUPERVISED LEARNING DALAM PENGOLAHAN DATA. *Jurnal Teknologi Terapan: G-Tech*, 4(2). <https://doi.org/10.33379/gtech.v4i2.635>
- Agustian, N., and Salsabila, U. H. (2021). Peran Teknologi Pendidikan dalam Pembelajaran. *ISLAMIKA*, 3(1). <https://doi.org/10.36088/islamika.v3i1.1047>

- Andriansyah, D. (2022). Implementasi Extract-Transform-Load (ETL) Data Warehouse Laporan Harian Pool. *Jurnal Teknik Informatika*, 8(2). <https://doi.org/10.51998/jti.v8i2.486>
- Berliantara, A. Y., Wicaksono, S. A., and Pinandito, A. (2017). Optimasi Scheduling untuk Proses Extract, Transform, Load (ETL) pada Data Warehouse Menggunakan Metode Round Robin Data Partitioning (Studi Kasus: Universitas XYZ). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(11).
- Effendi, D., and Wahidy, D. A. (2019). Pemanfaatan Teknologi Dalam Proses Pembelajaran Menuju Pembelajaran Abad 21. *Prosiding Seminar Nasional Pendidikan Program Pascasarjana Universitas Pgri Palembang*.
- Nofitri, R., and Irawati, N. (2019). ANALISIS DATA HASIL KEUNTUNGAN MENGGUNAKAN SOFTWARE RAPIDMINER. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 5(2). <https://doi.org/10.33330/jurteksi.v5i2.365>
- Oslan, Y., and Kristanto, D. H. (2019). Proses ETL (Extract Transformation Loading) Data Warehouse Untuk Peningkatan Kinerja Biodata Dalam Menyajikan Profil Mahasiswa Dari Dimensi Asal Sekolah Studi Kasus: Biodata Mahasiswa UKDW. In *Research Fair Unisri* (Vol. 3, Issue 1).
- Putra, I. M. S., and Adhitya Putra, D. K. T. (2019). Rancang Bangun Engine ETL Data Warehouse dengan Menggunakan Bahasa Python. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(2). <https://doi.org/10.29207/resti.v3i2.872>
- Retnoningsih, E., and Pramudita, R. (2020). Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python. *BINA INSANI ICT JOURNAL*, 7(2). <https://doi.org/10.51211/biict.v7i2.1422>
- Wijaya, R., and Pudjoatmodjo, B. (2016). Penerapan Extraction-Transformation-Loading (ETL) Dalam Data Warehouse (Studi Kasus : Departemen Pertanian). *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 5(2). <https://doi.org/10.23887/janapati.v5i2.9855>

Halaman ini sengaja dikosongkan