**Analisis Kepadatan Pekerjaan Pengolahan Data di Indonesia: Studi Kasus Scraping LinkedIn**

Ardiansyah Indra Febrianto1, Bagas Cahya Fajar Bastian2

Jurusan Sains Data Terapan, Politeknik Elektronika Negeri Surabaya,Institut Teknologi Sepuluh Nopember, Jl. Raya ITS, Keputih, Kec. Sukolilo, Surabaya, Jawa Timur 60111[1][2]

indrardi92@gmail.com1, bagascahyafajarbastian@gmail.com2

**Abstrak**

**Penelitian ini bertujuan untuk menganalisis distribusi geografis kepadatan pekerjaan di bidang pengolahan data di Indonesia berdasarkan data yang diambil dari platform LinkedIn. Dengan menggunakan metode web scraping, kami mengumpulkan informasi mengenai lowongan pekerjaan terkait pengolahan data seperti Data Analyst, Data Scientist, dan Data Engineer. Data ini kemudian dianalisis untuk mengidentifikasi wilayah dengan konsentrasi tertinggi dalam menyediakan pekerjaan di bidang ini, serta perusahaan yang mendominasi pasar kerja. Selain itu, penelitian ini juga mengeksplorasi keterampilan dan kualifikasi yang paling sering diminta oleh perusahaan di berbagai kota besar di Indonesia. Hasil penelitian ini diharapkan memberikan wawasan mengenai distribusi peluang karir dalam bidang pengolahan data, sekaligus membantu para profesional dan pencari kerja dalam memahami dinamika pasar kerja di berbagai wilayah.**

***Kata Kunci* – Scraping, LinkedIn, Data Analyst, Data Science, Data Engineer**

1. Pendahuluan

Berkembangnya teknologi digital telah meningkatkan kebutuhan akan ahli dalam manajemen dan analisis data. Kemampuan untuk menggunakan data dalam pengambilan keputusan strategis menjadi salah satu keterampilan yang semakin dicari di berbagai industri. Pekerja di bidang pengolahan data, seperti analis data, ilmuwan data, dan insinyur data, sangat penting untuk mengatasi kesulitan dalam mengelola volume besar data yang dibuat setiap hari.

Permintaan tenaga kerja di bidang ini di Indonesia terus meningkat, terutama seiring dengan pertumbuhan bisnis berbasis teknologi dan transformasi digital di berbagai sektor. Tetapi pekerjaan di bidang pengolahan data ini tidak merata di seluruh Indonesia. Kota-kota besar seperti Jakarta, Bandung, dan Surabaya seringkali memiliki banyak pekerjaan, sementara daerah lain mungkin memiliki lebih sedikit peluang.

Tujuan dari penelitian ini adalah untuk mengumpulkan data lowongan kerja pengolahan data di LinkedIn untuk menggambarkan bagaimana pekerjaan pengolahan data tersebar di seluruh Indonesia. Penelitian ini diharapkan dapat membantu pencari kerja dalam merencanakan karir mereka dan menemukan peluang kerja yang tepat dengan memahami wilayah mana yang memiliki banyak peluang kerja. Selain itu, tujuan penelitian ini adalah untuk mengidentifikasi keterampilan dan kualifikasi yang paling dibutuhkan oleh perusahaan dan memberikan gambaran tentang tren kompetensi di pasar kerja saat ini.

1. Tinjauan Pustaka

**2. 1 Web Scraping**

Web scraping adalah proses pengumpulan data web terstruktur secara otomatis yang dilakukan oleh aplikasi khusus atau kode pemrograman.[1]

Web Scraping dapat dilakukan baik secara manual oleh seorang pengguna atau secara otomatis oleh bot atau crawler web. Proses Web Scraping dari Internet dapat dibagi menjadi dua langkah berurutan, yaitu mengakuisisi sumber daya web dan kemudian mengekstraksi informasi yang diinginkan dari data yang diperoleh. [2]

Manfaat dari web scraping adalah agar informasi yang dikeruk lebih terfokus sehingga memudahkan dalam melakukan pencarian sesuatu.[3]

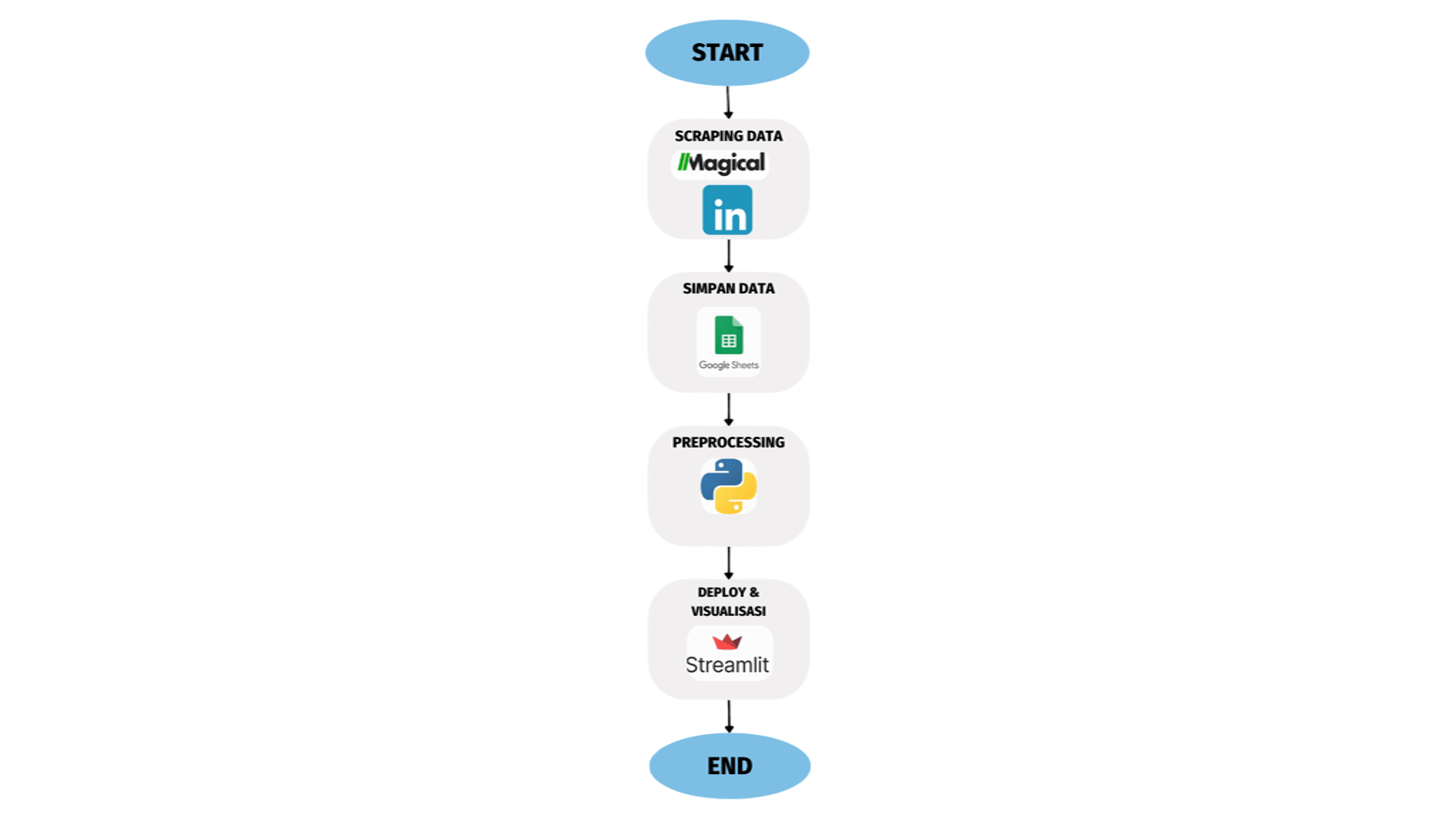
**2.2 LinkedIn**

LinkedIn adalah platform online untuk menampung pencapaian profesional[4]. Linkedin didirikan pada desember 2002 akan diluncurkan pada tanggal 5 Mei 2003[5].

Linkedin merupakan platform media social professional dengan jumlah 443 juta pengguna di sekitar 200 negara di dunia[6]. Pengguna LinkedIn di Indonesia didominasi oleh generasi milenial dengan rentang usia berkisar antara 18 hingga 24 tahun[7]. Misi Linkedin adalah untuk mendukung perusahaan yang membutuhkan tenaga kerja dan perekrutan (Mendapatkan bakat dan merekrut kumpulan bakat)[8].

1. Methodologi

Methodologi yang digunakan dalam perancangan dan implementasi dari *web scraping* ini antara lain:



Gambar 1. Use Case Diagram.

Berikut alur kerja dari “Analisis Kepadatan Pekerjaan Pengolahan Data di Indonesia: Studi Kasus Scraping LinkedIn”

* Scraping Data dilakukan dengan menggunakan *Extension* Magical melalui platform LinkedIn. Data diambil dari *filterisasi* pekerjaan melalui keyword “Data Science”.
* Data hasil scraping diletakkan pada spreadsheet sebagai tempat pengumpulan data.
* Dataset yang telah terkumpul diunduh dan dilakukan preprocessing dengan menggunakan Python untuk menghapus variabel dan kata yang tidak diperlukan pada pengolahan berikutnya.
* Dataset yang telah bersih digunakan untuk menampilkan peta dengan menambahkan titik koordinat letak kota/ provinsi/ negara yang membuka lowongan pekerjaan tersebut.
* Hasil visualisasi akan digunakan sebagai data tambahan yang bisa dibaca pengguna melalui deployment yang diletakkan pada platform streamlit.

*3.1 Scraping LinkedIn*

Dalam penelitian ini, data dikumpulkan melalui metode web scraping yang dilakukan pada platform pencarian kerja LinkedIn untuk memperoleh informasi terkait pekerjaan mengenai “Data Science” seperti Lokasi, Nama Perusahaan, Tipe Pekerjaan dan Jam Kerja. Proses scraping dilakukan dengan menggunakan Magical untuk mengekstraksi data dari halaman web secara otomatis. Data yang berhasil dikumpulkan kemudian diolah dan dianalisis menggunakan perangkat lunak **pandas** untuk mengidentifikasi tren dalam profil pengguna, pola pekerjaan, dan keterampilan yang relevan.

*3.2 Desain Tabel*

Setelah proses scraping selesai, data yang dikumpulkan disimpan dalam format spreadsheet. Dataset ini terdiri dari 15 kolom dan 265 baris yang memuat berbagai informasi penting mengenai lowongan yang ditemukan. Menyimpan data dalam spreadsheet memudahkan dalam melakukan analisis lebih dalam dan mengatur informasi secara lebih terstruktur.

|  |  |  |
| --- | --- | --- |
| **No** | **Nama Kolom** | **Keterangan** |
| 1 | Job Title | Posisi yang ditawarkan |
| 2 | Company Name | Nama Perusahaan |
| 3 | Location | Banyak pelamar pada Perusahaan |
| 4 | URL | Tautan deskripsi pekerjaan |
| 5 | Company Location | Lokasi Perusahaan |
| 6 | Company Size | Jumlah karyawan Perusahaan |
| 7 | Industry Stats | Data statistik profil Perusahaan |
| 8 | Job Qualifications | Kualifikasi akademik relevan |
| 9 | Job Skills | Kemampuan yang diperlukan |
| 10 | Job Skills | Kemampuan yang diperlukan |
| 11 | Job Skills | Kemampuan yang diperlukan |
| 12 | Profile | Profil perusahaan |
| 13 | Skills | Kemampuan yang diperlukan |
| 14 | Type | Tipe pekerjaan |
| 15 | Work Hours | Jam Kerja |

Tabel 1. Data Scraping

*3.3 Preprocessing Data*

Setelah pengumpulan data, langkah selanjutnya adalah pra pemrosesan untuk memastikan kualitas dan kesesuaian data untuk digunakan dalam analisis. Kumpulan data yang disimpan dalam spreadsheet diunduh dan kemudian diproses dengan Python. Proses ini melibatkan penghapusan variabel dan kata-kata yang tidak perlu, membuat kumpulan data bersih yang siap untuk analisis lebih lanjut. Pembersihan data ini penting untuk menghindari bias dan memastikan analisis yang dilakukan akurat dan juga dapat diandalkan.

*3.4 Deployment*

Dataset yang telah melalui proses pembersihan kemudian digunakan untuk membuat visualisasi peta yang menunjukkan lokasi kota, provinsi atau negara yang memiliki lowongan di bidang pengolahan data. Hasil visualisasi ini akan disajikan secara interaktif melalui platform Streamlit sehingga memungkinkan pengguna untuk mengeksplorasi data dengan lebih mudah dan intuitif. Oleh karena itu, penelitian ini tidak hanya memberikan informasi mengenai kepadatan kerja tetapi juga memungkinkan pencari kerja merencanakan karirnya berdasarkan visualisasi yang jelas dan informatif.

Referensi

[1] Yondra, A. S., Triyanto, D., & Bahri, S. (2022). Implementasi Web Scraping Untuk Mengumpulkan Informasi Produk Dari Situs E-Commerce dan Marketplace Dengan Teknik Pemrosesan Paralel. Coding : Jurnal Komputer Dan Aplikasi, 10(1), 93–102.

[2] Flores, V. A., Permatasari, P. A., & Jasa, L. (2020). Penerapan web scraping sebagai media pencarian dan menyimpan artikel ilmiah secara otomatis berdasarkan keyword. Maj. Ilm. Teknol. Elektro, 19(2), 157.

[3] Yani, D. D. A., Pratiwi, H. S., & Muhardi, H. (2019). Implementasi web scraping untuk pengambilan data pada situs marketplace. JUSTIN (Jurnal Sistem dan Teknologi Informasi), 7(4), 257-262.

[4] LinkedIn, O. (2022). About linkedin. LinkedIn Corporation.

[5] Martono, A., Maulani, G., & Pujianingsih, S. (2017). Pengembangan Web Alumni dengan Menggunakan Linkedin pada Perguruan Tinggi. Cyberpreneurship Innovative and Creative Exact and Social Science, 3(1), 71-83.

[6] Putra, H. M. M., & Fahamsyah, M. H. (2021). Penerapan Platform Media Sosial Linkedin Sebagai Alat Rekrutmen. Jurnal Investasi, 7(4), 15-24.

[7] Indriyana, S., Muthmainah, T. N., & Ridha, A. A. (2023). ANALISIS USER EXPERIENCE PADA PLATFORM LINKEDIN MENGGUNAKAN METODE USER EXPERIENCE QUESTIONNAIRE. METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi, 7(2), 208-216.

[8] Jauhary, H. Perbandingan metode analisis sentimen support vector machine, naïve bayes, dan logistic regression (Studi kasus: ULASAN GOOGLE PLAYSTORE APLIKASI LINKEDIN) (Bachelor's thesis, Fakultas Sains dan Teknologi UIN Syarif HIdayatullah Jakarta).

[9] Adila, N. (2022). Implementation of web scraping for journal data collection on the SINTA website. Sinkron: jurnal dan penelitian teknik informatika, 6(4), 2478-2485.

[10] Flores, V. A., Permatasari, P. A., & Jasa, L. (2020). Penerapan web scraping sebagai media pencarian dan menyimpan artikel ilmiah secara otomatis berdasarkan keyword. Maj. Ilm. Teknol. Elektro, 19(2), 157.