

ANALISIS DATA PENCARI KERJA DI DINAS PERINDUSTRIAN TENAGA KERJA KOTA SURABAYA DENGAN MENGGUNAKAN EDA DAN PCA

Triyatul Dewi Safitri, Amalia Anjani Arifiyanti, Agung Brastama Putra

Sistem Informasi, UPN “Veteran” Jawa Timur
Jalan Rungkut Madya Surabaya, Indonesia
triyatuldewisafitri@gmail.com

ABSTRAK

Ketidakseimbangan dalam penyaluran tenaga kerja di Kota Surabaya menjadi tantangan bagi Dinas Perindustrian dan Tenaga Kerja Kota Surabaya. Pencari kerja mempunyai latar belakang yang beragam seperti tingkat pendidikan, usia pengalaman kerja dan kondisi lainnya yang dapat mempengaruhi kesiapan dari pencari kerja. Penelitian ini bertujuan untuk menganalisis data pencari kerja di Dinas Perindustrian dan Tenaga Kerja Kota Surabaya menggunakan *Exploratory Data Analysis* (EDA) dan *Principal Component Analysis* (PCA). EDA digunakan untuk mengidentifikasi pola dan karakteristik pencari kerja berdasarkan faktor – faktor seperti usia, tingkat pendidikan, pengalaman kerja, status pernikahan, status disabilitas, dan lainnya. PCA diterapkan untuk mereduksi dimensi data serta menemukan faktor utama yang berkontribusi terhadap kesiapan kerja. Hasil penelitian ini menunjukkan bahwa reduksi dimensi menghasilkan PC 1, PC2 dan PC 3 yang menjelaskan total varians terbesar dalam data. Faktor dominan yang mempengaruhi pengelompokan pencari kerja adalah terdapat 3 komponen utama yang mempunyai korelasi yang cukup besar terhadap pembentukan pengelompokan pencari kerja adalah nama program studi dengan nilai varians paling tinggi pertama yaitu 97,2%, diikuti oleh kombinasi faktor usia, pengalaman kerja, status pernikahan, jumlah anak dan status disabilitas dengan varians 1,51% serta faktor pengalaman kerja, tingkat pendidikan, sertifikat mendapatkan varians 1,04%.

Kata kunci : pencari kerja, *Exploratory Data Analysis* (EDA), *Principal Component Analysis* (PCA), Surabaya

1. PENDAHULUAN

Di Provinsi Jawa Timur jumlah pencari kerja setiap tahun mengalami peningkatan dengan angka mencapai 200.388 orang pada tahun 2023 [1].

Pencari kerja tersebut berasal dari berbagai latar belakang yang berbeda – beda, hal ini dapat mempengaruhi ketidakseimbangan dalam penyaluran tenaga kerja khususnya di Kota Surabaya. Di Kota Surabaya, Dinas Perindustrian dan Tenaga Kerja melalui *website* ASSIK (Arek Suroboyo Siap Kerjo) berperan sebagai perantara untuk memberikan informasi lowongan pekerjaan kepada masyarakat serta membantu perusahaan dalam menemukan kandidat yang sesuai dengan kualifikasi yang dibutuhkan.

Berdasarkan hal tersebut, perlu dilakukan analisis data pencari kerja di Kota Surabaya berdasarkan karakteristik yang sama dengan melakukan pemahaman data dan menemukan pola dalam data pencari kerja menggunakan EDA (*Exploratory Data Analysis*) dan PCA (*Principal Component Analysis*).

PCA digunakan untuk mengurangi jumlah dimensi dari kumpulan data, sehingga lebih mudah untuk dipahami namun, tetap mempertahankan informasi penting pada data [2].

Dengan menerapkan PCA, jumlah awal variabel sebanyak n akan dikurangi menjadi k yaitu variabel baru yang disebut dengan *principal components* [3].

Beberapa penelitian sebelumnya juga menggunakan EDA dan PCA yang berjudul “Analisis Pengaruh Faktor pada Stunting di Jawa Timur Tahun 2021 Menggunakan Metode *Principal Component*

Analysis (PCA)” penelitian ini menganalisis faktor kesejahteraan dan kesehatan masyarakat dan berhasil mengidentifikasi 2 faktor utama yang dapat menjelaskan 69,4888% dari variansi total dalam data, hasil ini menunjukkan bahwa dari kedua faktor utama merupakan komponen utama yang mempengaruhi stunting di Jawa Timur [4].

Penelitian yang dilakukan oleh [3] menyatakan bahwa metode PCA efektif untuk reduksi dimensi data dan dapat mengidentifikasi faktor utama yang mempengaruhi data.

Selanjutnya penelitian EDA yang dilakukan oleh [5] bertujuan untuk memberikan informasi serta *knowledge* untuk melakukan pengambilan keputusan dengan lebih cepat. Penelitian ini menggabungkan kedua teknik analisis EDA dan PCA dengan harapan memperoleh wawasan yang lebih mendalam terkait karakteristik pencari kerja di Kota Surabaya. Dengan hasil analisis ini diharapkan mampu mendukung pengambilan keputusan yang lebih tepat dalam mengatasi ketidakseimbangan pencari kerja di Surabaya.

2. TINJAUAN PUSTAKA

2.1. Pencari Kerja

Pencari kerja merupakan sekumpulan individu yang aktif mencari pekerjaan untuk mencari pendapatan, jenjang karir atau pengembangan diri.

Menurut penelitian [6] pencari kerja merupakan angkatan kerja yang mempunyai status aktif dalam mendaftarkan diri pada informasi lowongan kerja. Pencari kerja terbentuk dari beberapa faktor yaitu dari

kondisi ekonomi, demografi, atau preferensi individu terhadap bidang pekerjaan yang diinginkan.

Penelitian sebelumnya telah menerapkan EDA untuk memahami distribusi data. Studi dalam [7] membahas penerapan EDA pada kualitas susu sapi, yang menunjukkan bahwa metode ini dapat mengidentifikasi pola distribusi data. Studi dalam [3] menerapkan PCA dalam analisis pengidap kanker serviks, di mana PCA berhasil mereduksi dimensi data dan menentukan faktor dominan yang mempengaruhi kondisi pasien.

Berdasarkan penelitian tersebut, metode EDA dan PCA dapat diterapkan untuk menganalisis pola pencari kerja serta mengidentifikasi faktor utama yang mempengaruhi kesiapan kerja.

2.2. EDA

Exploratory Data Analysis (EDA) merupakan tahapan untuk pemahaman data dengan menganalisis data dengan menggunakan berbagai pendekatan, *outputnya* untuk meringkas serta menggambarkan fitur utama dari kumpulan data [8].

EDA bertujuan untuk mencari pola berdasarkan dataset untuk menemukan anomali data, memahami korelasi dari beberapa variabel [7].

2.3. PCA

Principal Component Analysis (PCA) merupakan analisis komponen utama dari teknik yang digunakan untuk menyederhanakan data sehingga terbentuk sistem koordinat baru dengan varian maksimal tanpa kehilangan karakteristik penting dari data tersebut secara signifikan [9].

Metode analisis PCA dapat efektif jika data yang ada mempunyai jumlah variabel yang besar dan mempunyai korelasi antar masing – masing variabel selanjutnya untuk perhitungan dari PCA berdasarkan nilai eigen serta vektor eigen yang memberikan pernyataan terkait persebaran data dari suatu dataset [3].

3. METODE PENELITIAN

Metode penelitian ini yaitu menggunakan metode pendekatan *exploratory data analysis (EDA)* dan *principal component analysis (PCA)* yang merupakan tahapan pengumpulan data, melakukan identifikasi pada data sampai menemukan pola pada data dan menarik informasi atau makna yang dapat digunakan untuk pengambilan keputusan. Tahapan yang dilakukan pada proses penelitian yaitu sebagai berikut:

3.1. Analisis Masalah

Permasalahan utama dalam penelitian ini adalah sulitnya memahami distribusi dan karakteristik pencari kerja berdasarkan berbagai faktor seperti usia, tingkat pendidikan, pengalaman kerja, status pernikahan dan faktor lainnya. Ketidakseimbangan dalam data pencari kerja dapat menghambat pengambilan keputusan dalam memberikan rekomendasi pelatihan atau

lowongan pekerjaan yang sesuai dengan kebutuhan industri.

3.2. Data Acquisition

Pada penelitian yang digunakan dataset adalah data pencari kerja yang mendaftar melalui *website ASSIK (Arek Suroboyo Siap Kerjo)*. Didalam dataset terdapat 9 variabel yang mencakup informasi seperti tingkat pendidikan, usia, nama prodi, pengalaman kerja, sertifikat, status pernikahan, jumlah anak dan status disabilitas dan jumlah total dataset yang digunakan dalam penelitian ini berjumlah 2100 pencari kerja. Data tersebut kemudian diproses lebih lanjut untuk memastikan kualitasnya sebelum dilakukan analisis.

3.3. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) adalah salah satu teknik analisis yang digunakan pada penelitian ini dengan fungsi untuk melakukan pengecekan dan persiapan sebagai pengujian hipotesis, melakukan penanganan pada nilai data yang hilang (*missing values*), melakukan pengecekan pada tipe data dan melakukan transformasi pada data kategorik menjadi data numerik menggunakan label encoding selanjutnya melihat visualisasi dari grafik distribusi usia, tingkat pendidikan, pengalaman kerja serta hubungan antar variabel untuk lebih memahami pola dalam dataset penelitian.

3.4. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) pada penelitian ini menggunakan *cumulative variance ratio* untuk menentukan dari jumlah komponen utama yang dipertahankan, selanjutnya dapat dilihat persentase varians total yang dapat dijelaskan dari setiap komponen utama.

Komponen utama yang dilakukan oleh penelitian sebelumnya mencapai 70 – 95% dari varians total. *Ratio* merupakan komponen utama hingga mencapai *cumulative variance* yang cukup besar seperti dari PC1 dan 2 PC2 menjelaskan 80% varians total [10].

4. HASIL DAN PEMBAHASAN

Hasil dari proses analisis data yang telah dilakukan menggunakan *Exploratory Data Analysis (EDA)* dan *Principal Component Analysis (PCA)* akan ditampilkan dalam bentuk visualisasi tabel, grafik untuk mempermudah pemahaman terhadap pola yang ditemukan.

4.1. Analisis Masalah

Ketidakseimbangan penyaluran tenaga kerja perlu dilakukan analisis lebih dalam agar menghasilkan informasi yang terstruktur serta bermakna. Informasi ini digunakan untuk mengidentifikasi pola yang terdapat pada data pencari kerja di Kota Surabaya, sehingga dapat memberikan wawasan yang lebih mendalam.

4.2. Data Acquisition

Proses pengambilan data diimplementasikan pencari kerja didapatkan dari Dinas Tenaga Kerja Kota Surabaya, untuk melakukan persiapan dataset dapat diimplementasikan untuk membaca dataset dengan *library Pandas* dan membaca dataset dan dataset pencari kerja yang mendaftar melalui *website ASSIK* (Arek Suroboyo Siap Kerjo) Kota Surabaya diubah menggunakan format CSV dan diakses melalui *path google drive* untuk memudahkan pengolahan data.

Selanjutnya dilakukan pengecekan duplikasi pada data untuk memastikan bahwa tidak ada input data yang tercatat lebih dari satu kali. Proses ini bertujuan untuk menjaga kualitas data agar analisis yang dilakukan tidak bias atau terpengaruh oleh data yang berulang. Jika ditemukan data duplikat, maka akan dilakukan proses penghapusan agar dataset tetap akurat dan representatif.

Tabel 1. Pengecekan Duplikasi Data

Jumlah Data yang terduplikasi	0
--------------------------------------	---

Pada tabel 1 dilakukan pengecekan duplikasi data namun, tidak ada data yang terduplikasi sehingga dapat dilakukan pada langkah selanjutnya yaitu melakukan pengecekan pada *missing value* dataset pencari kerja yang mendaftar melalui *website ASSIK* (Arek Suroboyo Siap Kerjo) Kota Surabaya.

Tabel 2. Tabel Pengecekan Missing Value

No	0
Tanggal	0
Nama	0
Nama_pendidikan	0
Usia	0
Nama_prodi	55
Pengalaman Kerja	0
Sertifikat	0
Disabilitas	0
Status Pernikahan	5
Jumlah_anak	0

Pada tabel 2 dilakukan pengecekan *missing value* terdapat nilai yang kosong pada atribut *nama_prodi* dan *status_pernikahan*, untuk penanganannya akan dilakukan pengisian *missing value* menggunakan *mean* dan untuk hasil penanganan *missing value* dapat dilihat pada tabel 3.

Tabel 3. Hasil Penanganan Missing Value

No	0
Tanggal	0
Nama	0
Nama_pendidikan	0
Usia	0
Nama_prodi	0
Pengalaman Kerja	0
Sertifikat	0
Disabilitas	0
Status Pernikahan	0
Jumlah_anak	0

Selanjutnya Pada tabel 4 dilakukan tahap *selection data* yang dimana data yang tidak relevan untuk penelitian pencari kerja sehingga perlu untuk dihapus. Atribut yang dihapus pada penelitian ini adalah atribut *tanggal*, *nama* dan *no*.

Tabel 4. Penghapusan Atribut

Atribut yang dihapus	
No	Dihapus
Tanggal	Dihapus
Nama	Dihapus

Setelah atribut data asli ada yang sebagian dihilangkan dikarenakan tidak mengandung informasi yang relevan. Dilanjutkan pada tahap pengecekan tipe data untuk memastikan bahwa setiap atribut mempunyai format yang sesuai dengan kebutuhan analisis. Jika ditemukan ketidaksesuaian, maka dilakukan proses konversi tipe data agar lebih sesuai, seperti mengubah nilai numerik yang tersimpan sebagai teks kategorikal menjadi tipe numerik. Hasil dari pengubahan tipe data ini akan diproses menjadi atribut yang digunakan pada penelitian, seperti yang terlihat pada tabel 5.

Tabel 5. Atribut yang Digunakan

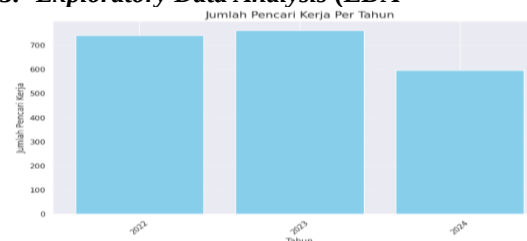
Tingkat_pendidikan	int
Usia	int
Nama_prodi	int
Pengalaman Kerja	int
Sertifikat	int
Disabilitas	int
Sertifikat	int
Status_pernikahan	int
Jumlah_anak	int

Selanjutnya pada tabel 6 dilakukan transformasi menggunakan *label encoding*, isi atribut yang di transformasi yaitu tingkat pendidikan, nama prodi, pengalaman kerja, sertifikat, disabilitas, status pernikahan.

Tabel 6. Data 5 Teratas Transformasi Data Kategori menjadi Numerik

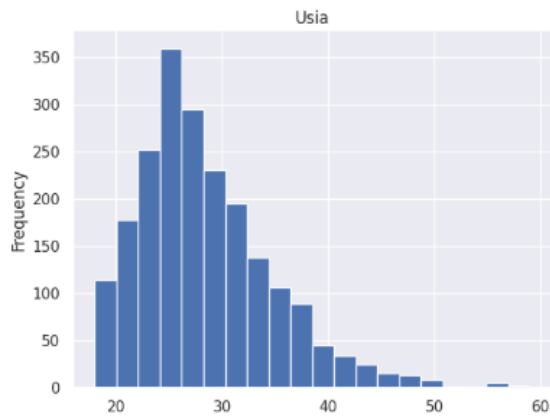
TP	US	NP	PK	SRT	DS	SP	JA
7	20	121	0	1	0	0	0
8	30	41	0	1	1	0	0
7	25	138	0	1	0	0	0
2	42	62	0	1	0	3	1
8	26	135	0	1	0	0	0

4.3. Exploratory Data Analysis (EDA)



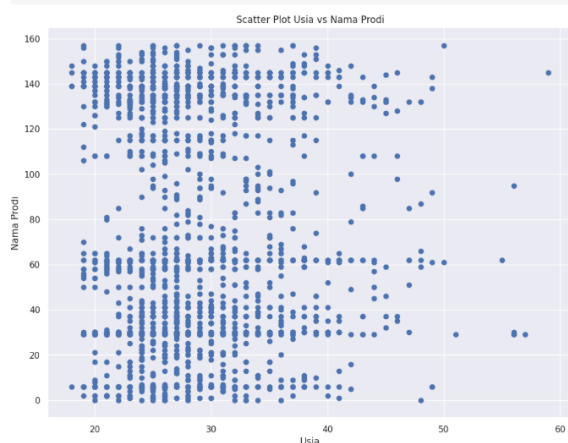
Gambar 1. Diagram Batang Pencari Kerja

Grafik pada gambar 1 tersebut menunjukkan jumlah pencari kerja per tahun dalam dataset yang digunakan. Dari grafik diagram batang, terlihat adanya fluktuasi jumlah pencari kerja. Pada tahun 2022, jumlah pencari kerja yang relatif tinggi, tahun 2023 mengalami sedikit peningkatan pada tahun 2022, yang menunjukkan adanya peningkatan jumlah pencari kerja yang terdaftar. Namun, pada tahun 2024 terjadi penurunan jumlah pencari kerja dibandingkan dengan tahun sebelumnya. Analisis lebih lanjut diperlukan untuk memahami penyebab perubahan dan dampaknya terhadap kebijakan ketanagakerjaan.



Gambar 2. Histogram Pada Atribut Usia

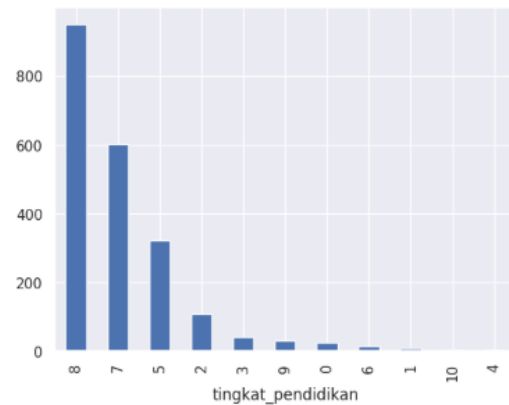
Distribusi pencari kerja yang terlihat pada gambar 2 menunjukkan bahwa mayoritas berasal dari rentang usia dari 20 hingga 30 tahun, dengan puncak frekuensi pada usia 20 ke atas. Hal ini mengindikasikan bahwa sebagian besar pencari kerja mempunyai pengalaman kerja, sedang mencari peningkatan *karier* atau sedang mengalami transisi pekerjaan.



Gambar 3. Scatter Plot Usia dengan Nama Prodi

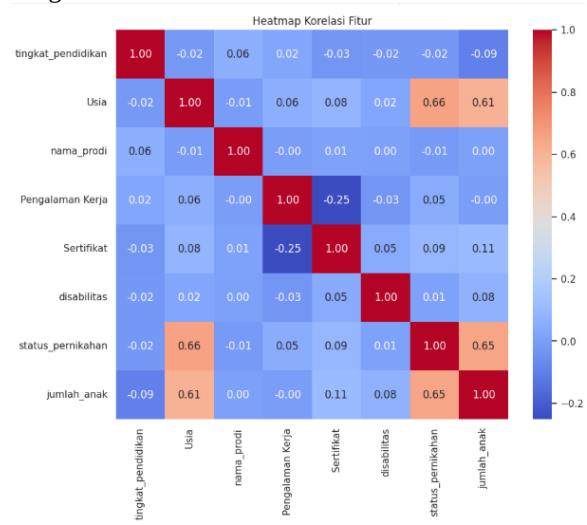
Pada gambar 3 yang memvisualisasikan *Scatter plot* antara atribut usia dengan nama prodi (program studi) menunjukkan sebaran pencari kerja berdasarkan usia dengan latar belakang pendidikan masing – masing pencari kerja. Dari visualiasi *scatter plot* dapat dianalisis bahwa sebagian besar pencari kerja berasal

dari rentang 20 hingga 30 tahun, dengan distribusi yang bervariasi di setiap prodi.



Gambar 4. Diagram Batang Atribut Tingkat Pendidikan

Pada gambar 4 Diagram batang tingkat pendidikan menunjukkan distribusi pencari kerja berdasarkan jenjang pendidikan yang sudah ditempuh. Dari grafik tersebut terlihat bahwa mayoritas pencari kerja berasal dari tingkat pendidikan pada kategori S1 yaitu di angka 8. Distribusi yang semakin menurun menunjukkan bahwa jumlah pencari kerja dengan pendidikan rendah relatif sedikit dibandingkan dengan pencari kerja yang mempunyai pendidikan yang lebih tinggi. Hal tersebut dapat mengindikasikan bahwa pencari kerja dengan pendidikan rendah mempunyai keterbatasan dalam mengakses informasi atau kurangnya keterampilan yang sesuai dengan pasar kerja dan sebaliknya tingginya jumlah pencari kerja dari tingkat pendidikan yang lebih tinggi bisa mencerminkan persaingan ketat di dunia kerja atau kurangnya kesesuaian antara keterampilan lulusan dengan kebutuhan industri.

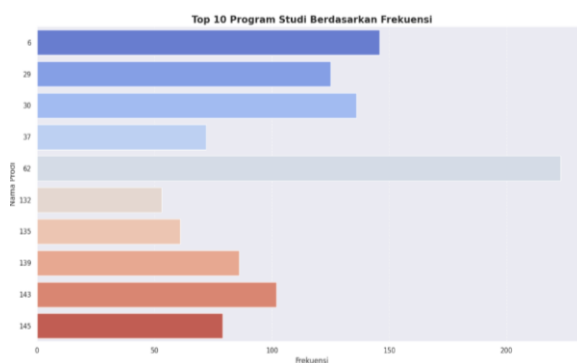


Gambar 5. Heatmap Korelasi Fitur

Berdasarkan *heatmap* korelasi fitur pada gambar 5 terdapat beberapa hubungan menarik yang dapat dianalisis yaitu dari atribut usia dengan status

pernikahan terdapat korelasi positif yang cukup kuat dengan tingkat keterkaitan di angka 0.66, hal ini dapat dianalisis bahwa pertambahan usia dapat menyebabkan jumlah pernikahan meningkat. Selanjutnya korelasi positif antara atribut usia dengan jumlah anak di angkat 0.61 yang menunjukkan bahwa semakin bertambah usia, umumnya pencari kerja mempunyai lebih banyak anak. Selanjutnya korelasi tinggi didapatkan dari atribut status pernikahan dengan jumlah anak dengan angka 0.65 dapat diartikan bahwa pencari kerja yang sudah berstatus menikah cenderung mempunyai anak.

Sebaliknya terdapat korelasi yang cenderung negatif lemah antara usia dengan tingkat pendidikan dapat dianalisis bahwa tingkat pendidikan tidak selalu berbanding lurus dengan bertambahnya usia karena banyak faktor seperti putus sekolah atau menempuh pendidikan non – formal dan korelasi negatif di angka -0,25 dari atribut pengalaman kerja dengan sertifikat yang menunjukkan bahwa sertifikasi tidak selalu didapat oleh pencari kerja yang mempunyai pengalaman kerja karena bisa jadi sertifikat lebih banyak diperoleh oleh pencari kerja yang baru memasuki dunia kerja untuk meningkatkan peluang dalam mendapatkan pekerjaan.

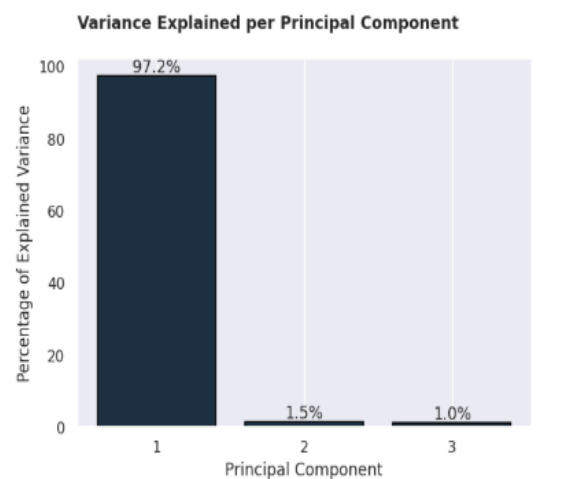
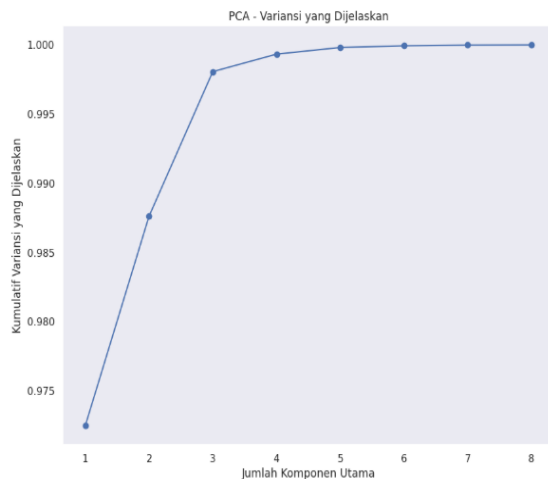


Gambar 6. Diagram Batang Frekuensi Nama Prodi

Selanjutnya pada gambar 6 distribusi pencari kerja berdasarkan frekuensi nama program studi dapat dianalisis bahwa beberapa program studi dengan frekuensi tertinggi hal ini menunjukkan bahwa jurusan – jurusan tersebut mempunyai banyak lulusan yang sedang mencari pekerjaan dan beberapa jurusan mempunyai pencari kerja yang relatif sedikit yang dapat mengindikasikan bahwa lulusan dari jurusan tersebut lebih mudah mendapatkan pekerjaan atau jumlah lulusannya memang sedikit.

Program studi dengan jumlah pencari kerja tertinggi menunjukkan kesulitan dalam mendapatkan pekerjaan, bisa jadi terdapat ketidakseimbangan antara jumlah lulusan dengan kebutuhan pasar kerja.

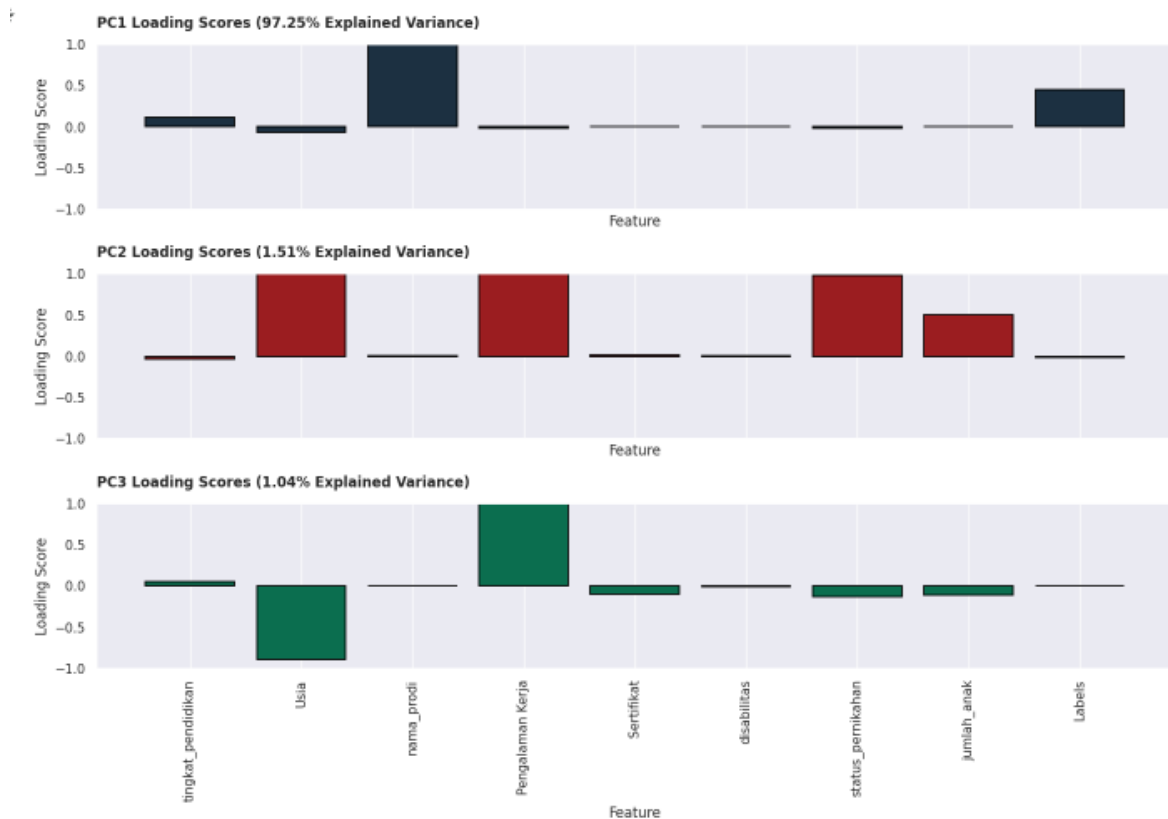
4.4. Principal Component Analysis (PCA)



Gambar 7. Variansi Kumulatif yang Dijelaskan PCA

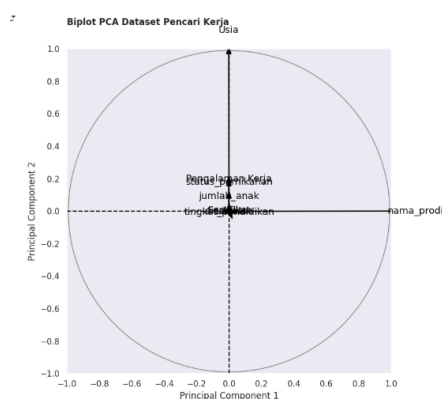
Pada gambar 7 Varians kumulatif adalah total persentase varians yang mampu dijelaskan dari komponen utama 1 hingga komponen utama yang lain. Jumlah komponen utama akan dipilih untuk menentukan yang paling optimal berdasarkan kumulatif varians.

Pada komponen utama 1 (PC 1) mencakup informasi sebesar 97% yang mempunyai makna komponen ini sudah mencakup mayoritas informasi pada data. Pada grafik terlihat bahwa jumlah komponen utama yang mampu mencakup informasi sebesar 99% dari komponen utama 1 (PC 1) hingga komponen utama 3 (PC 3) dari total varians.



Gambar 8. Plot Skor Pemuatan PCA

Pada gambar 8 dari plot skor pemuatan (*loading scores*) yang ditampilkan, dapat dianalisis peran masing – masing terhadap komponen utama (*principal components* / PC). PC1 menjelaskan hampir seluruh variasi dalam data yang mencapai 97,25% yang artinya nama_prodi merupakan faktor utama dalam menentukan pola kesiapan kerja. Selain nama prodi terdapat beberapa atribut lain dengan *loading score* yang cukup tinggi yaitu atribut usia, tingkat pendidikan, pengalaman kerja.



Gambar 9. Biplot PCA

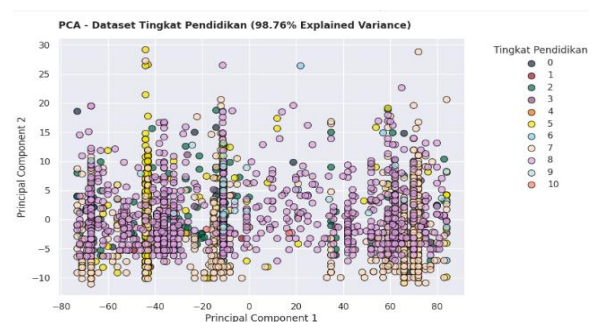
Selanjutnya PC 2 (1,51%) atribut usia, status pernikahan dengan jumlah anak mempunyai pengaruh kecil terhadap kesiapan kerja dan PC 3 (1,04%) Sertifikasi dan pengalaman kerja lebih dominan yang

artinya sertifikasi dapat menjadi faktor tambahan dalam kesiapan kerja tetapi tidak sebesar nama prodi.

Pada gambar 9 PC 1 pada biplot PCA Dataset Pencari Kerja adalah atribut nama prodi karena mempunyai korelasi tinggi dengan PC 1 dengan vektornya hampir sejajar dengan sumbu X yang menunjukkan bahwa program studi merupakan faktor utama yang menjelaskan variasi dalam data.

Pada PC 2 adalah atribut usia yang mempunyai korelasi tinggi dengan PC2, karena vektornya hampir sejajar dengan sumbu Y yang menunjukkan bahwa usia merupakan faktor yang membedakan data pada dimensi kedua.

Sedangkan vektor dari atribut ditengah dan pendek menunjukkan bahwa pengaruhnya terhadap dua komponen utama ini tidak sebesar nama prodi atau usia, namun keberadaannya tetap penting dalam analisis keseluruhan.



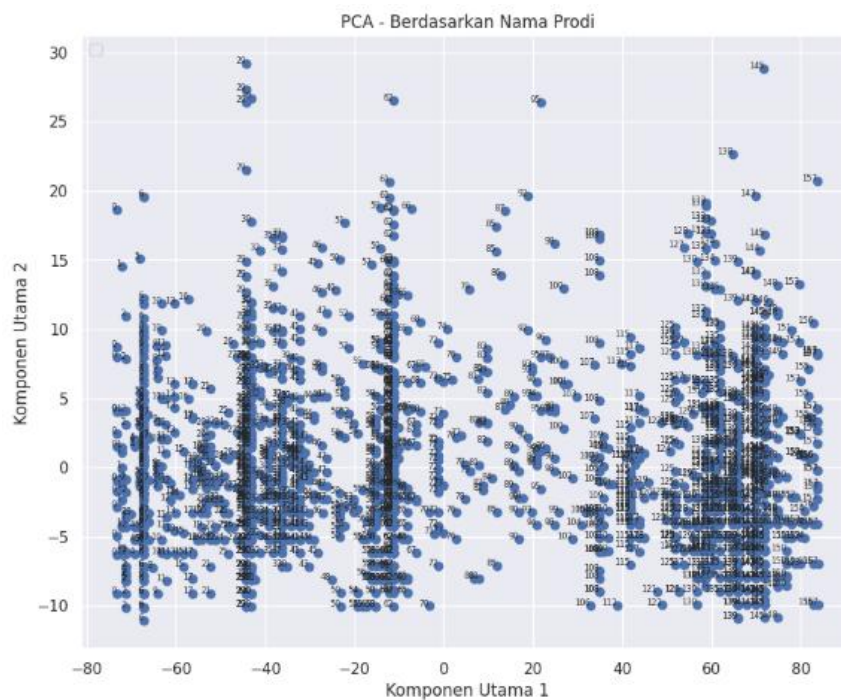
Gambar 10. Plot Varians Tingkat Pendidikan

Pada gambar 10 dari plot PCA yang telah ditampilkan, dapat dilihat bahwa tingkat pendidikan mempunyai persebaran yang unik dalam ruang PCA dan plot pada grafik diatas menampilkan kode tingkat pendidikan. Adapun *insight* yang dapat diambil adalah pendidikan tinggi S1,S2,Diploma IV dengan inisiasi kode 3,8,9 cenderung mempunyai penyebaran yang lebih luas pada *Principal Component 1* (PC1), yang artinya terdapat variasi yang signifikan dalam karakteristik pencari kerja seperti pencari kerja dengan pendidikan tinggi mempunyai latar belakang pengalaman kerja,sertifikasi yang lebih beragam.

Selanjutnya pada bagian tengah plot adalah dengan tingkat pendidikan SLTA/MA dan SMK/MAK dengan inisiasi kode 5 dan 7, menunjukkan bahwa pencari kerja mempunyai pola distribusi yang relatif beragam dibandingkan dengan kelompok pendidikan tinggi dengan perbedaan utama antara SLTA / MA dan SMK/MAK dapat terjadi pada aspek keterampilan dan kesiapan kerja langsung.

Pendidikan dasar dan tidak memiliki ijazah (SD/MI, SLTP/MTS, Tidak Memiliki Ijazah) dengan inisiasi kode 4,6,10 mempunyai persebaran yang lebih terbatas dan lebih terkonsentrasi di area tertentu dalam ruang PCA yang mempunyai makna bahwa kelompok pencari kerja ini mempunyai karakteristik yang lebih seragam seperti dari segi keterampilan dan kesiapan kerja.

Selanjutnya pada gambar 11 plot PCA yang telah ditampilkan, dapat dilihat bahwa nama prodi mempunyai persebaran dalam ruang PCA dan plot pada grafik diatas menampilkan kode tingkat pendidikan. Pada kode nama prodi 1 berada di sebelah kiri yang mewakili prodi non-teknik sedangkan kode 62 berada di tengah yang mewakili prodi campuran non teknik dan teknik sedangkan di sebelah kanan terdapat kode 132 yang mewakili prodi teknik dengan kondisi yang lebih padat sehingga banyak program studi teknik terkelompok dalam satu bagian, maka Kota Surabaya mempunyai pencari kerja dari bidang teknik yang membutuhkan pelatihan tambahan agar lebih sesuai dengan kebutuhan industri.



Gambar 11. Plot Varians Nama Prodi

5. KESIMPULAN DAN SARAN

Secara keseluruhan 8 atribut, 2 atribut menjadi faktor utama dalam mengukur peluang pencari kerja untuk mendapatkan pekerjaan yaitu atribut program studi dengan usia, sedangkan 6 atribut lainnya merupakan pengukuran yang berkaitan dengan faktor pelengkap dalam pengukuran peluang pencari kerja. Adapun PC1 menjelaskan 97,25% variasi dalam data, menunjukkan bahwa nama program studi mempunyai pengaruh terbesar dalam peluang pekerjaan, hal ini mengindikasikan bahwa latar belakang pendidikan sangat menentukan dalam mendapatkan pekerjaan

sehingga pencari kerja dari program studi yang relevan dengan kebutuhan industri mempunyai peluang lebih besar untuk siap bekerja dibandingkan pencari kerja yang latar belakang pendidikannya tidak selinear dengan pasar kerja.

Selanjutnya PC 2 menjelaskan 1,51% menunjukkan bahwa usia yang mempunyai dominasi lebih tinggi selanjutnya diikuti oleh pengalaman kerja, status pernikahan dan jumlah anak mempunyai pengaruh kecil terhadap pengukuran peluang dalam mendapatkan pekerjaan yang artinya meskipun aspek ini berpengaruh namun bukan mereka faktor utama

yang menentukan peluang dalam mendapatkan pekerjaan, hal ini mengindikasikan bahwa pencari kerja dari berbagai usia dan kondisi sosial tetap mempunyai peluang yang sama tergantung pada kompetensi utama pencari kerja dan pengalaman kerja pada PC 2 ini menunjukkan pencari kerja yang lebih tua atau sudah menikah sehingga mempunyai pengalaman kerja namun tidak sesuai dengan bidang pekerjaan.

PC 3 menjelaskan 1,04% yang melihat kontribusi atau peran sertifikasi dan pengalaman kerja sebagai faktor pelengkap dikarenakan sertifikasi dapat meningkatkan kesiapan kerja, namun dampaknya tidak sebesar nama prodi dan pengalaman kerja juga menjadi faktor penting, namun hanya sebagai pelengkap dari pendidikan formal dan sertifikasi yang dipunyai oleh pencari kerja.

Berdasarkan hasil analisis, dapat disimpulkan bahwa faktor utama yang menentukan peluang pencari kerja untuk mendapatkan pekerjaan adalah nama program studi, yang mencerminkan relevansi pendidikan dengan kebutuhan industri. Sementara faktor pengalaman kerja dan sertifikasi berperan sebagai faktor pendukung yang dapat meningkatkan kesiapan kerja, tetapi tidak sebesar pengaruh program studi. Faktor – faktor lain seperti usia, status pernikahan, jumlah anak dan status disabilitas mempunyai pengaruh namun tidak sebesar nama program studi.

Sebagai saran, pencari kerja disarankan untuk meningkatkan daya saing dengan memperoleh keterampilan dan sertifikasi tambahan yang sesuai dengan bidang industri yang diminati. Selain itu, Dinas Tenaga Kerja dan Perindustrian khususnya di Kota Surabaya dapat lebih mengerahkan program pengembangan keterampilan yang relevan dengan industri, sehingga pencari kerja dari berbagai latar belakang pendidikan dan beragam kondisi dapat meningkatkan peluang untuk mendapatkan pekerjaan.

Dengan pendekatan yang lebih strategis dalam penyesuaian pelatihan dan kebutuhan industri diharapkan kesenjangan antara pencari kerja dan dunia kerja dapat dikurangi, sehingga meningkatkan tingkat kesiapan kerja dan peluang mendapatkan pekerjaan yang sesuai dengan kebutuhan dan kondisi masyarakat Kota Surabaya.

DAFTAR PUSTAKA

- [1] BPS, “Pencari Kerja Terdaftar, Lowongan Kerja Terdaftar, dan Penempatan/Pemenuhan Tenaga Kerja Menurut Provinsi dan Jenis Kelamin, 2023,” *Badan Pusat Statistik*, 2023. <https://www.bps.go.id/id/statistics-table/3/VEU5VVVERIVWM0JwYTNvdk1ISkpWR3R1VUhaVmR6MDkjMw==/pencari-kerja-terdaftar-lowongan-kerja-terdaftar-dan-penempatan-pemenuhan-tenaga-kerja-menurut-provinsi-dan-jenis-kelamin.html?year=2023> (accessed Sep. 21, 2024).
- [2] S. Manullang, N. Kairani, M. S. Sinaga, and B. Hutapea, “ANALISIS FAKTOR PENYEBAB PENYAKIT JANTUNG MENGGUNAKAN METODE PRINCIPAL COMPONENT ANALYSIS (PCA) masyarakat . Penyakit jantung juga merupakan masalah kesehatan yang kritis karena terikat , melainkan mencari saling ketergantungan antar variabel untuk meng,” vol. 5, no. 3, 2024.
- [3] M. Z. Nasution, A. A. Nababan, K. U. Syaliman, M. S. Novelan, and M. Jannah, “Penerapan Principal Component Analysis (PCA) Dalam Penentuan Faktor Dominan Yang Mempengaruhi Pengidap Kanker Serviks (Studi Kasus : Cervical Cancer Dataset),” *J. Mantik Penusa*, vol. 3, no. 1, pp. 204–210, 2019.
- [4] K. Kunci, “Analisis Pengaruh Faktor pada Stunting di Jawa Timur Tahun 2021 Menggunakan Metode Principal Component Analysis (PCA) Dzakki Damar Rabbani Institut Teknologi Sepuluh Nopember , Indonesia PENDAHULUAN Stunting adalah kondisi yang mencerminkan masalah gizi,” vol. 2, no. 11, pp. 1–13, 2024.
- [5] A. Wibowo, “Analisa Dan Visualisasi Data Penjualan Menggunakan Exploratory Data Analysis Pada PT. Telkominfra,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 3, pp. 2292–2304, 2022, doi: 10.35957/jatisi.v9i3.2737.
- [6] Irwansya, Sahrul Ramadhan, and Teguh Ansyor Lorosae, “Perancangan Sistem Informasi Pencari Kerja Pada Dinas Tenaga Kerja Dan Transmigrasi Kabupaten Kep. Sula Propinsi Maluku Utara Menggunakan Visual Basic.Net,” *J. Mediat.*, vol. 7, no. 2, pp. 107–113, 2024, doi: 10.59562/mediatik.v7i2.1304.
- [7] R. Ramadhani, R. Ramadhanu, and T. Hidayat, “Exploratory Data Analysis (EDA) untuk Mengetahui Distribusi Data Kualitas Susu Sapi,” *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 23, no. 1, p. 68, 2024, doi: 10.53513/jis.v23i1.9500.
- [8] D. T. Husni *et al.*, “Analisis Big Data Penjualan Video Games Menggunakan Eda,” *J. Tek. Inf. dan Komput.*, vol. 5, no. 1, p. 43, 2022, doi: 10.37600/tekinkom.v5i1.517.
- [9] R. Fikri, A. Mushardiyanto, M. N. Laudza’Banin, K. Maureen, and H. Patria, “Pengelompokan Kabupaten/Kota di Indonesia Berdasarkan Informasi Kemiskinan Tahun 2020 Menggunakan Metode K-Means Clustering Analysis,” *Semin. Nas. Tek. dan Manaj. Ind.*, vol. 1, no. 1, pp. 190–199, 2021, doi: 10.28932/sentekmi2021.v1i1.76.
- [10] S. P. Wulandari, N. P. Sari, R. Aprili, and G. Imani, “Analysis of Indonesian Public Health Factors 2023 with PCA Method and Factor Analysis,” no. 1, pp. 1–15, 2025.