

# OPTIMASI DATA MINING AKADEMIK MENGUNAKAN ALGORITMA PARTICLE SWARM OPTIMIZATION DAN KNN UNTUK KLASIFIKASI TOPIK TUGAS AKHIR MAHASISWA

*By* Sumarni Sumarni

1

# OPTIMASI DATA MINING AKADEMIK MENGGUNAKAN ALGORITMA PARTICLE SWARM OPTIMIZATION DAN KNN UNTUK KLASIFIKASI TOPIK TUGAS AKHIR MAHASISWA

7

Sumarni<sup>a,1</sup>, Suhardi Rustam<sup>a,2</sup><sup>a</sup>Universitas Ichsan Gorontalo, Jln. Achmad Najamuddin, Kota Gorontalo dan Kode Pos 96115, Indonesia<sup>1</sup>marnisiwa@gmail.com <sup>2</sup>suhardirstm@gmail.com

| INFORMASI ARTIKEL                                                                                                                                                                                             | ABSTRAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Diterima : xx – xx – 2020<br/>Direvisi : xx – xx – 2020<br/>Diterbitkan : xx – xx – 2020</p> <p><b>Kata Kunci:</b><br/>Topik Tugas Akhir<br/>Klasifikasi<br/>KNN<br/>PSO (Particle Swarm Optimization)</p> | <p>Permasalahan Topik Tugas akhir merupakan suatu bentuk karya tulis ilmiah yang memuat hasil pengamatan dari suatu penelitian terhadap masalah yang terjadi dengan menggunakan metode yang berkaitan dengan bidang ilmu tertentu. Setiap mahasiswa disetiap program studi harus menyusun tugas akhir. Namun, sebelum memulai menulis tugas akhir, setiap mahasiswa harus mempunyai bidang topik sebagai tujuan, langkah pemilihan topik tugas akhir merupakan langkah awal sebelum mengerjakan tugas akhir. Salah satu cara untuk mendapatkan topik tugas akhir yaitu dengan melihat nilai mata kuliah umum serta mata kuliah konsentrasi jurusan, nilai yang mendominasi adalah nilai yang layak untuk cakupan topik penelitian. penelitian ini dilakukan penerapan metode K-Nearest Neighbor (KNN) untuk kategorisasi nilai mata kuliah konsentrasi untuk cakupan topik penelitian, topik seluruh nilai dalam dataset nantinya akan diklasifikasikan oleh KNN dan di optimasi dengan algoritma Particle swarm Optimization (PSO). Eksperimen kategorisasi tugas akhir ini dibangun dengan data latih Mahasiswa Universitas Ichsan Gorontalo yang telah diklasifikasikan sebelumnya dan data uji berasal dari seluruh nilai Mata Kuliah yang belum diketahui kategorinya. Hasil eksperimen, nilai akurasi yang dihasilkan algoritma KNN yaitu nilai akurasi terbaik dengan K=3, K Folds = 10 memiliki akurasi yaitu 72.46% dan Algoritma KNN-PSO akurasi terbaik dengan K=3, K Folds = 10 memiliki akurasi yaitu 89.86%, ini menunjukan akurasi lebih baik dengan menggunakan algoritma optimasi</p>                                                                                                   |
| <p><b>Keywords:</b><br/>Topic of the Final Assignment<br/>Classification<br/>KNN<br/>PSO (Particle Swarm Optimization)</p>                                                                                    | <p><b>ABSTRACT</b><br/>Problems the Topic of the final project is a form of scientific writing that contains the results of observations from a study of the problems that occur with the use of methods related to the particular field of science. Every student in every program of study must draw up a final project. However, before embarking on writing the final project, each student must have the topic area as a destination, the step of selection the topic of final project is an initial step before working on the final task. One way to get the final task is to see the value of general courses as well as courses, concentration majors, the value of which dominate the is is decet to scope the research topic. this research is conducted on the application of the method of K-Nearest Neighbor (KNN) for categorization of the value of the courses of concentration for the coverage of the research topic, topic the entire value in the dataset will be classified by KNN and in the optimization with the Particle swarm Optimization algorithm (PSO). The experimental categorization of the final project is built with the training data Mahasiswa Universitas Ichsan Gorontalo that has been classified previously and test data derived from the entire value of the courses is not yet known categories. The results of the experiments, the value of the resulting accuracy of algorithms KNN, namely the value of the best accuracy with K=3, K Folds = 10 has an accuracy that is 72.46% and the Algorithm of KNN-PSO best accuracy with K=3, K Folds = 10 has an accuracy that is 89.86%, shows the accuracy is better by using the optimization algorithm</p> |

## 1. PENDAHULUAN

Melimpahnya Data Akademik yang dimiliki Universitas sebagai aset informasi yang akan berguna dalam pengelolaan akademik dan sebagai *benefit value* yang akan menguntungkan universitas. Data tersebut diantaranya tentang data akademik mahasiswa. Di bidang akademik, disetiap semester bertambahnya jumlah data yang direkan data dari kegiatan akademik seperti perekaman tugas mahasiswa, nilai mid, tugas tambahan mahasiswa dan nilai

akhir semester mahasiswa dan sebagainya sehingga mengakibatkan tumpukan data dengan berbagai kapasitas, dengan meningkatnya volume penyimpanan pada situs dan server akademik universitas. Hal tersebut juga berlaku di Teknik Informatika Universitas Ichsan Gorontalo. TA(Tugas Akhir) merupakan bentuk penelitian yang dilakukan mahasiswa di akhir masa studi sebagai salah satu persyaratan untuk kelulusan mahasiswa tersebut. Namun tidak sedikit mahasiswa yang akan melaksanakan Tugas Akhir kesulitan dalam mencari topik Tugas Akhir yang sesuai dengan keahlian dan minat mahasiswa tersebut. Padahal ada banyak penelitian atau Tugas Akhir terdahulu yang dapat dikembangkan sebagai topik Tugas Akhir yang baru bagi mahasiswa. Dalam penelitian ini akan dirancang pengklasifikasian suatu dataset akademik yang dapat membantu mahasiswa-mahasiswa yang akan melaksanakan Tugas Akhir kuliah di Program Studi Teknik Informatika dan kemudian klasifikasi topik yang mungkin sesuai dengan minat dan keahlian dari mahasiswa tersebut. Rekomendasi dilakukan dengan hasil dari klasifikasi berdasarkan cakupan nilai akademik terdahulu yang dilakukan sebelumnya oleh mahasiswa Teknik Informatika Universitas Ichsan Gorontalo Dengan adanya data tersebut nantinya diharapkan dapat membantu mahasiswa yang akan melaksanakan Tugas Akhir untuk menemukan topik yang sesuai dengan keahlian dan minat mahasiswa, sehingga dapat mengurangi jumlah mahasiswa yang terlambat lulus dikarenakan sulit menemukan topik yang sesuai dengan minat dan keahlian mahasiswa.

Dataset akademik mahasiswa akan melalui proses pengolahan dari data menjadi pengetahuan ini disebut dengan istilah data mining. Data mining adalah disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau menemukan pola dari suatu data yang besar. Pengetahuan yang dihasilkan dapat berupa pola, rumus, aturan atau model. Training *data mining* mempelajari bagaimana mengolah data menjadi pengetahuan menggunakan *software* data mining orange. Peserta akan mendapatkan banyak studi kasus penerapan Data Mining. Diharapkan setelah mengikuti training ini, peserta siap menghadapi tantangan kasus-kasus pada penerapan *data mining* pada kehidupan nyata.

Sebagai algoritma optimasi digunakan Algoritma PSO (Particle Swarm Optimization) sebagai salah satu algoritma optimasi yang dapat digunakan untuk pengambilan keputusan. Tetapi bisa juga digunakan untuk pencarian jalur[1]. Contoh yang dibahas kali ini adalah mengenai pencarian posisi dengan pengembalian nilai fungsi minimal[2].

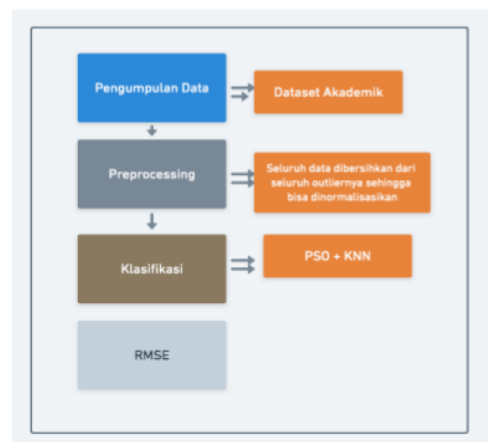
Particle Swarm Optimization[3] adalah teknik optimasi dengan cara menghitung secara terus menerus calon solusi dengan menggunakan suatu acuan kualitas. Adapun algoritma K-NN[4] merupakan metode untuk melakukan klasifikasi terhadap objek[5] berdasarkan data yang jaraknya paling dekat dengan objek tersebut[6], cara untuk mengukur jarak kedekatan antara data baru dengan data lama(data training)[7], diantaranya *Euclidian distance* [8]dengan menggunakan sampel-sampel dari data testing[9].

Pada penelitian akan menganalisis bagaimana mengoptimasi algoritma KNN dengan menggunakan algoritma PSO (Particle Swarm Optimization) dalam menentukan klasifikasi topik penelitian mahasiswa semester akhir dan bagaimana akurasi dari penggunaan optimasi dengan hasil optimal

## II. METODE

Metode eksperimen digunakan untuk penelitian ini, penggunaan cara penelitian eksperimen untuk tahap penelitian akan dimulai dari pengumpulan data sampai dengan mendapatkan hasil evaluasi dan hasil untuk mendapatkan tujuan yang akan dicapai, tahapan preprocessing sebagai tahap awal untuk mengidentifikasi *missing value* dan menghilangkan data yang tidak memiliki nilai[10], seperti yang tergambar pada model penelitian sebagai berikut :

### A. Model Eksperimen



Gambar 1. Model Ekperimen KNN-PSO Validation

## B. Pengumpulan Data

Identifikasi variable-variabel dari dataset di bagian pusat pengelolaan data akademik universitas sebagai tahap awal sebagai cara dalam menentukan pengelolaan data dan proses analisisnya menggunakan alat analisis data[11]. Beberapa Variable yaitu seluruh mata kuliah konsentrasi yang terkait, nilai dari setiap mata kuliah dan nilai prestasi akademik

## C. Pengolahan Data Awal

Pada tahapan ini melakukan *preprocessing* terhadap data untuk menghilangkan outlier[12], melengkapi nilai setiap variable yang terkait sehingga memiliki kesiapan dalam penggunaan alat analisis data

## D. Eksperimen

Eksperimen pada tahapan ini adalah menjalan seluruh prosedur yang telah ditentukan/disusun dalam sebuah kerangka pemikiran untuk menyelesaikan masalah[13], adapun tahapannya adalah, di mulai dari penyiapan dataset kemudian dilakukan langkah prepossessing terhadap data untuk menghilangkan outliernya[14], selanjutnya adalah menormalisasikan data kemudian menggunakan tool rapidminer data mining untuk mengelola data tersebut[15] dengan menggunakan metode *Optimasi dan KNN* pada tahapan optimasi dan klasifikasi yang hasilnya akan di evaluasi dengan akurasi yang teroptimasi.

## E. Evaluasi

Pada tahapan Pengujian eksperimen yang dilakukan pada algoritma KNN dan KNN-PSO serta eksperimen nilai akurasi pada algoritma optimasi tersebut[16].

# III. HASIL DAN PEMBAHASAN

## A. Deskripsi Data eksperimen

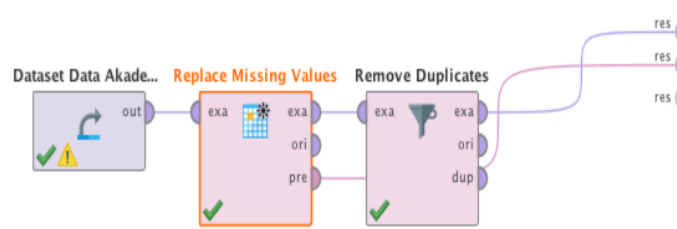
Deskripsi data eksperimen merupakan info data yang digunakan dalam eksperimen ini

Table 1. Info Data Eksperimen

| No | Nama Uraian                     | Deskripsi                                                                            |
|----|---------------------------------|--------------------------------------------------------------------------------------|
| 1  | Dataset Data Akademik Mahasiswa | Nama dataset yang berisi sekumpulan data yang siap digunakan dalam eksperimen        |
| 2  | Dataset Size                    | Ukuran dataset dalam bentuk kolom dan baris                                          |
| 3  | Features                        | Model data yang akan diproses seperti kategori dan numerik                           |
| 4  | Meta Atribut                    | Variable dalam dataset yang akan dilewati dalam pemrosesan data seperti nim dan nama |

## B. Model preprocessing

Model processing diawali pada tahapan preprocessing, dimana seluruh data harus menghilangkan outlier/missing value yang tidak memiliki nilai variable dalam dataset



### C. Data Hasil Preprocessing

Setelah melewati tahapan preprocessing, sebagian data tersebut ditampilkan sebagai berikut :

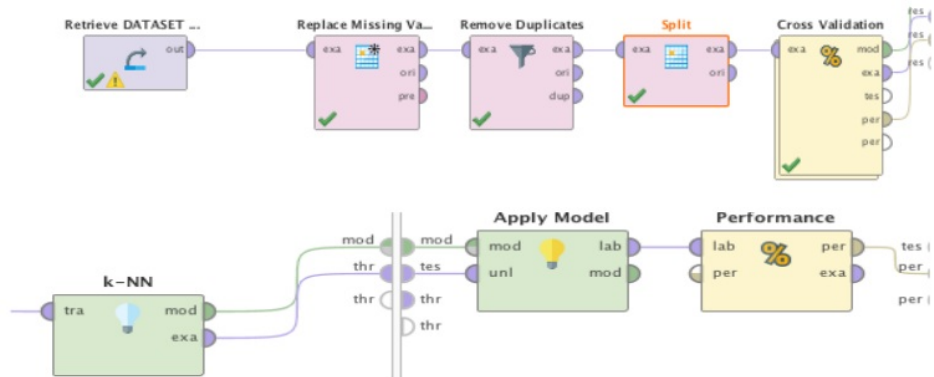
Tabel 2. Hasil Preprocessing

|    | A        | B             | C       | D          | E          | F          | G          | H          | I      | J         | K        | L        | M         | N        | O        | P      | Q        | R         |
|----|----------|---------------|---------|------------|------------|------------|------------|------------|--------|-----------|----------|----------|-----------|----------|----------|--------|----------|-----------|
|    | NIM      | Nama          | Kelamin | MKUII      | MKLOGIKOM  | MKELEARN   | MKEGOW     | MKRPL      | MKDATA | MKSTRUKTU | MKGARKOM | MKKOMLAN | MKDIGITAL | MKPROJCT | MKFRAMEW | MKMGNT | MKMOBILE | Label     |
| 1  | 73116032 | RANDI S BAI L |         | 0,13819444 | 4          | 0,252      | 0,10277778 | 0,36       | 0,07   | 0,06      | 0,06     | 4        | 0,07      | 0,06     | 0,06     | 4      | 0,07     | FRAMEWORK |
| 2  | 73116039 | GSKA PUTRI P  |         | 0,17152778 | 4          | 0,17638889 | 0,17152778 | 0,29       | 4      | 0,07      | 0,09     | 4        | 4         | 0,07     | 0,07     | 4      | 4        | FRAMEWORK |
| 3  | 73116030 | MUHAMMAD IL   |         | 0,17152778 | 0,06       | 0,17013889 | 0,17152778 | 4          | 4      | 0,07      | 0,07     | 0,06     | 4         | 0,07     | 0,07     | 0,06   | 4        | FRAMEWORK |
| 4  | 73116012 | DWI UMARU P   |         | 0,17708333 | 0,09       | 4          | 0,18611111 | 4          | 0,06   | 4         | 0,09     | 0,09     | 0,07      | 4        | 4        | 0,09   | 0,06     | FRAMEWORK |
| 5  | 73116014 | ISMAIL M S/L  |         | 0,25       | 0,07       | 0,17638889 | 0,18611111 | 0,54       | 4      | 4         | 4        | 0,07     | 4         | 4        | 4        | 0,07   | 4        | FRAMEWORK |
| 6  | 73116015 | AMIRIN PAPE L |         | 0,17152778 | 0,09       | 0,17638889 | 0,16875    | 0,17986111 | 4      | 0,07      | 0,07     | 0,09     | 4         | 0,07     | 0,07     | 0,09   | 4        | FRAMEWORK |
| 7  | 73116016 | NADIRA P      |         | 0,17708333 | 0,06       | 4          | 0,18611111 | 4          | 4      | 4         | 4        | 0,06     | 0,06      | 4        | 4        | 0,06   | 4        | FRAMEWORK |
| 8  | 73116017 | SARIN HUN Y   |         | 0,17152778 | 0,06       | 0,18541667 | 0,18611111 | 0,17430556 | 0,06   | 4         | 4        | 0,06     | 4         | 4        | 4        | 0,06   | 4        | FRAMEWORK |
| 9  | 73116018 | ALWIN Y LA P  |         | 0,17708333 | 0,07       | 0,17013889 | 0,18611111 | 0,54       | 0,09   | 0,06      | 0,06     | 0,07     | 4         | 0,06     | 0,06     | 0,07   | 0,09     | FRAMEWORK |
| 10 | 73116019 | NOVIYANTI P   |         | 0,34       | 0,07       | 0,18541667 | 0,17708333 | 4          | 0,07   | 4         | 4        | 0,07     | 4         | 4        | 4        | 0,07   | 4        | FRAMEWORK |
| 11 | 73116020 | ELA M PATU P  |         | 0,17708333 | 4          | 0,17013889 | 0,17708333 | 0,54       | 0,09   | 4         | 4        | 0,06     | 4         | 4        | 4        | 0,06   | 4        | FRAMEWORK |
| 12 | 73116021 | AKRIAL MU L   | 0       | 4          | 0,17916667 | 0,17708333 | 0,25       | 0,06       | 4      | 4         | 4        | 0,09     | 4         | 4        | 4        | 0,06   | 4        | FRAMEWORK |
| 13 | 73116023 | RENALDI TUI L |         | 0,17152778 | 0,07       | 0,17916667 | 0,16875    | 0,42       | 0,06   | 0,06      | 0,06     | 0,07     | 0,07      | 0,06     | 0,06     | 0,07   | 0,06     | FRAMEWORK |
| 14 | 73116024 | SUSANTO PU L  |         | 0,17708333 | 4          | 0,18541667 | 0,17708333 | 4          | 0,07   | 0,09      | 0,09     | 4        | 0,09      | 0,09     | 0,09     | 4      | 0,07     | FRAMEWORK |
| 15 | 73116025 | RAHMAT MP L   |         | 0,25       | 4          | 0,17638889 | 0,154      | 0,05       | 0,07   | 0,07      | 0,07     | 0,07     | 0,06      | 0,07     | 0,07     | 4      | 0,07     | FRAMEWORK |
| 16 | 73116026 | FIKRAM YUA L  |         | 0,33       | 0,06       | 0,17013889 | 0,17       | 0,21       | 4      | 0,09      | 0,09     | 4        | 0,06      | 0,09     | 0,09     | 0,06   | 4        | FRAMEWORK |
| 17 | 73116027 | MOH AFFAN L   |         | 0,17708333 | 4          | 0,17013889 | 0,17708333 | 0,17986111 | 4      | 4         | 4        | 4        | 0,07      | 4        | 4        | 4      | 4        | FRAMEWORK |
| 18 | 73116028 | WAHYU ARI L   |         | 0,14444444 | 4          | 0,17013889 | 0,17986111 | 0,07       | 0,07   | 4         | 4        | 0,07     | 0,07      | 0,07     | 0,07     | 4      | 0,07     | FRAMEWORK |
| 19 | 73116029 | NELVA PAUP P  |         | 0,17708333 | 4          | 0,17708333 | 0,17986111 | 4          | 4      | 0,06      | 4        | 4        | 4         | 4        | 4        | 4      | 4        | FRAMEWORK |
| 20 | 73116030 | DELVIANA DI P |         | 0,04       | 0,06       | 0,16786111 | 0,42       | 0,42       | 4      | 4         | 0,09     | 4        | 4         | 4        | 4        | 0,06   | 4        | FRAMEWORK |
| 21 | 73116032 | YATU M AB-P   |         | 0,17708333 | 0,09       | 0,18541667 | 0,17708333 | 0,42       | 0,06   | 0,06      | 0,07     | 0,06     | 0,07      | 0,06     | 0,06     | 0,09   | 0,06     | FRAMEWORK |
| 22 | 73116033 | MAINUN R P    |         | 0,17708333 | 0,07       | 0,18541667 | 0,18611111 | 4          | 4      | 4         | 0,09     | 4        | 4         | 4        | 4        | 0,07   | 4        | FRAMEWORK |
| 23 | 73116035 | ALAN MUSA L   |         | 0,25       | 0,09       | 0,17013889 | 0,154      | 0,08       | 4      | 4         | 4        | 0,06     | 4         | 4        | 4        | 0,09   | 4        | FRAMEWORK |
| 24 | 73116038 | FEBRI H NET L |         | 0,17708333 | 4          | 0,17013889 | 0,17708333 | 0,46       | 4      | 4         | 0,06     | 4        | 0,06      | 4        | 4        | 4      | 4        | FRAMEWORK |
| 25 | 73116039 | TRIHADRIYAN P |         | 0,16875    | 0,06       | 4          | 0,17708333 | 0,16875    | 0,06   | 0,06      | 0,07     | 0,06     | 4         | 0,06     | 0,06     | 0,06   | 0,06     | FRAMEWORK |
| 26 | 73116059 | ANDRA HUR L   |         | 0,54       | 0,09       | 0,17638889 | 0,17708333 | 4          | 0,09   | 0,09      | 0,07     | 0,09     | 4         | 0,09     | 0,09     | 0,09   | 0,09     | FRAMEWORK |
| 27 | 73116060 | FANDY UMA L   |         | 0,25       | 0,07       | 0,17916667 | 0,154      | 0,04       | 0,07   | 0,07      | 0,07     | 0,07     | 0,07      | 0,07     | 0,07     | 0,07   | 0,07     | FRAMEWORK |
| 28 | 73116061 | DIAN PRATIN P |         | 0,17708333 | 0,07       | 0,18541667 | 0,18611111 | 4          | 0,09   | 0,09      | 4        | 0,09     | 0,06      | 0,09     | 0,09     | 0,07   | 0,09     | EARNING   |
| 29 | 73116063 | ANSAR L       |         | 0,17708333 | 4          | 0,17638889 | 0,18611111 | 4          | 0,06   | 4         | 0,07     | 0,06     | 0,09      | 4        | 4        | 4      | 0,06     | EARNING   |
| 30 | 73116064 | RIZAL TUI L   |         | 0,23       | 4          | 0,17013889 | 0,17708333 | 3          | 4      | 4         | 4        | 0,06     | 0,07      | 4        | 4        | 4      | 4        | EARNING   |

### D. Algoritma KNN Validasi

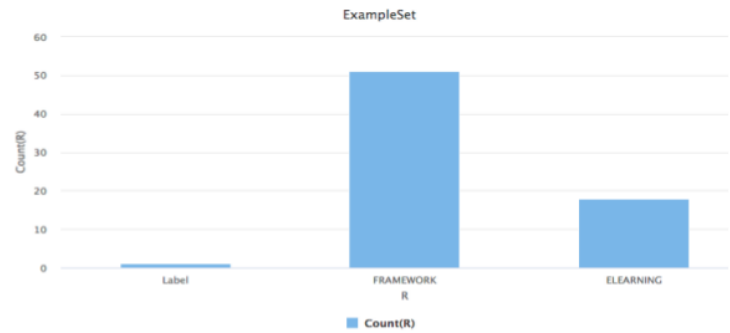
#### a. Model KNN Validasi eksperimen

Dalam model eksperimen ini algoritma KNN ditempatkan setelah data preprocessing, adapun gambarannya adalah sebagai berikut :



Gambar 3. Model KNN Validasi

Dalam memproses dataset dengan menggunakan model seperti pada gambar 4.3 yaitu menggunakan model KNN menghasilkan Klasifikasi yaitu Topik Framework dan Topik Elearning dengan grafik frekuensi kluster pada gambar 3.3 di bawah ini



Gambar 4. Grafik Frekuensi Klasifikasi

Hasil eksperimen data akademik menggunakan algoritma KNN Validasi adalah sebagai berikut

Table 3. Hasil Eksperimen dengan KNN Validation

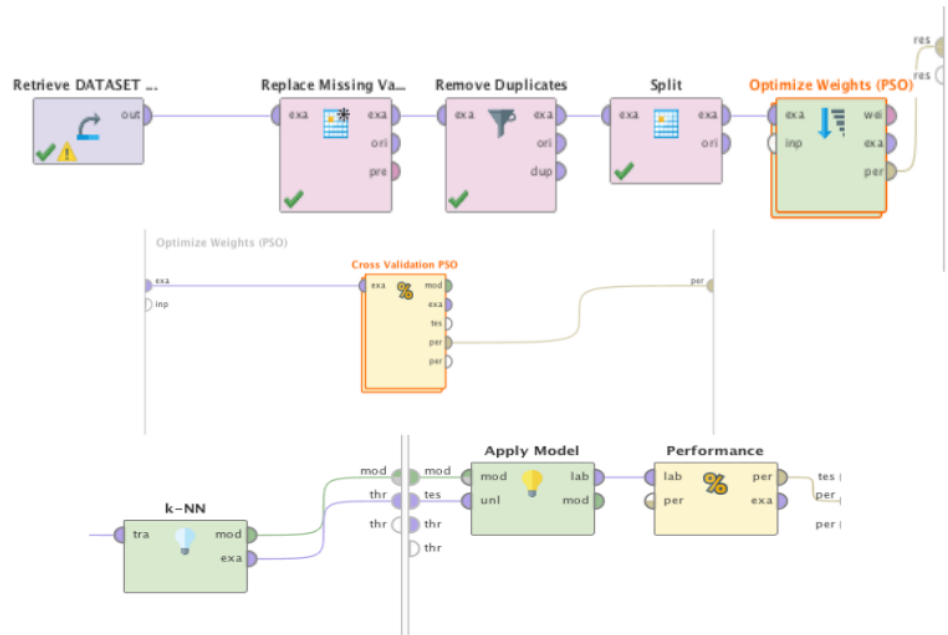
| No | Algoritma | Kriteria |         | Akurasi |
|----|-----------|----------|---------|---------|
|    |           | K        | K Folds |         |
|    | KNN       | 3        | 10      | 72.46%  |
|    |           | 4        | 10      | 72.46%  |
|    |           | 5        | 10      | 71.02%  |

Hasil eksperimen KNN Validation dengan kriteria adalah untuk K=3, K Folds = 10, Akurasinya = 72.46%, untuk K=4, K Folds = 10, Akurasinya = 72.46%, dan untuk K=5, K Folds = 10 memiliki Akurasi 71.01%, sehingga dari table 4.2 akurasi terbaik dengan K=3, K Folds = 10 memiliki akurasi yaitu 72.46%

## b. Algoritma KNN-PSO Validasi

### 1). Model Eksperimen

Pada model eksperimen ini menggunakan dataset dari data hasil KNN Preprocessing yang selanjutnya menjadi dataset untuk eksperimen dengan menggunakan KNN dengan optimasi PSO Validation dengan model sebagai berikut



Gambar 5. Model Eksperimen KNN dengan Optimasi PSO

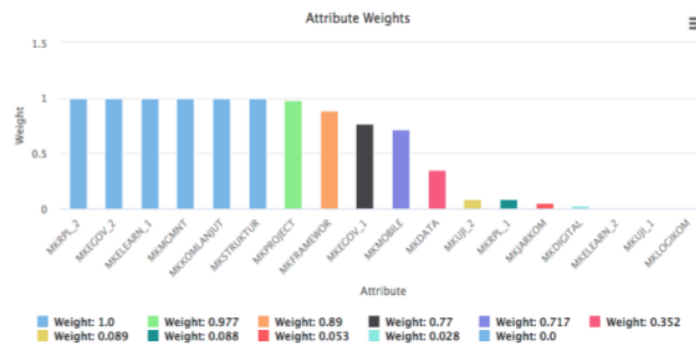
Eksperimen ini data yang digunakan adalah data dari hasil KNN Preprocessing dengan Model KNN Validation yang ditambahkan kedalam Operator Algoritma PSO sehingga KNN dalam model gambar 3.4 akan dioptimasi dengan algoritma PSO Validation, dengan hasil sebagai berikut :

Tabel 4. Data Hasil KNN-PSO Validation

| No | Algoritma | Kriteria |         | Akurasi |
|----|-----------|----------|---------|---------|
| 1. | KNN-PSO   | K        | K Folds |         |
|    |           | 3        | 10      | 89.86%  |
|    |           | 4        | 10      | 86.96%  |
|    |           | 5        | 10      | 71.01%  |

Hasil eksperimen KNN-PSO Validation dengan kriteria adalah untuk K=3, K Folds = 10, Akurasinya = 89.86%, untuk K=4, K Folds = 10, Akurasinya = 86.96%, dan untuk K=5, K Folds = 10 memiliki Akurasi 71.01%, sehingga dari table 4.3 akurasi terbaik dengan K=3, K Folds = 10 memiliki akurasi yaitu 89.86%

## 2). Distribusi data hasil KNN dengan K=2



Gambar 6. Grafik Optimize Weights (PSO)

Gambar 3.5, sebagai gambaran data dari model Optimize Weights (PSO) adalah bobot yang dimiliki setiap atribut setiap dataset pada Akademik Mata Kuliah dalam menentukan klasifikasi Topik Penelitian Mahasiswa semester akhir

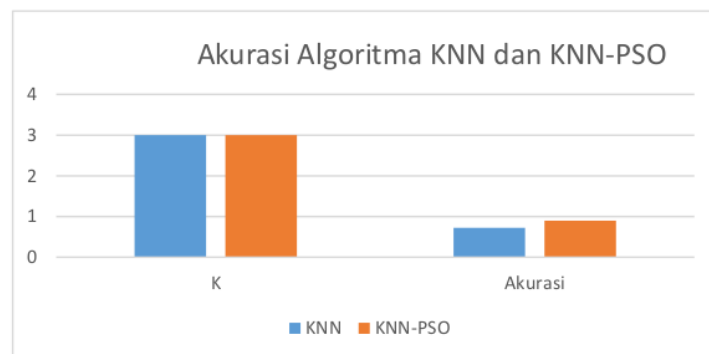
## 3). Akurasi data hasil Kmeans KNN dan KNN-PSO Validasi dimana K=3

Model klasifikasi dengan KNN Validation dan KNN-PSO Validation untuk K=3, K Folds = 10 dengan hasil eksperimen sebagai berikut :

Tabel 5. Data Hasil KNN dan KNN-PSO Validation

| No | Algoritma          | Kriteria |         | Akurasi |
|----|--------------------|----------|---------|---------|
|    | KNN Validation     | K        | K Folds |         |
|    |                    | 3        | 10      | 72.48%  |
|    |                    |          |         |         |
|    | KNN-PSO Validation | 3        | 10      | 89.86%  |

Adapun dalam bentuk grafik adalah sebagai berikut :



Gambar 7. Grafik Akurasi Algoritma KNN dan KNN-PSO

#### E. Pembahasan hasil eksperimen KNN Validation dan KNN-PSO Validation, K=3

Hasil eksperimen yang telah dilakukan dengan menggunakan dataset akademik mahasiswa dengan 70 record data, 14 atribut dan 1 label digunakan untuk menentukan topik penelitian yang diawali dengan Teknik preprosesing untuk membersihkan *outlier/missing value* pada setiap atribut yang ada dalam dataset. Pada eksperimen ini, tahapan pemodelan dengan menggunakan algoritma KNN sebagai tahapan pertama, tahapan kedua dengan menggunakan Algoritma PSO (Particle Swarm Optimization) untuk mengoptimasi model tahap awal. Pemodelan dengan menggunakan algoritma KNN Validasi dan juga pemodelan optimasi yaitu KNN-PSO Validasi.

Hasil eksperimen KNN Validation dengan kriteria adalah untuk K=3, K Folds = 10, Akurasinya = 72.46%, untuk K=4, K Folds = 10, Akurasinya = 72.46%, dan untuk K=5, K Folds = 10 memiliki Akurasi 71.01%, sehingga dari table 4.2 akurasi terbaik dengan K=3, K Folds = 10 memiliki akurasi yaitu 72.46%.

Pemodelan tahap kedua yaitu hasil eksperimen KNN-PSO Validation dengan kriteria adalah untuk K=3, K Folds = 10, Akurasinya = 89.86%, untuk K=4, K Folds = 10, Akurasinya = 86.96%, dan untuk K=5, K Folds = 10 memiliki Akurasi 71.01%, sehingga dari table 4.3 akurasi terbaik dengan K=3, K Folds = 10 memiliki akurasi yaitu 89.86%.

#### IV. Kesimpulan dan Saran

Hasil eksperimen dalam uraian pembahasan penelitian ini, dapat disimpulkan bahwa penerapan algoritma KNN dan Algoritma KNN-PSO, PSO (*Particle Swarm Optimization*) sebagai algoritma optimasi terhadap algoritma KNN, dimana K=3 menghasilkan klasifikasi topik penelitian mahasiswa semester akhir dan masing-masing klasifikasi di validasi dan memiliki nilai akurasi prediksi untuk kedua pemodelan validasi tersebut dan nilai Akurasi yang dihasilkan Algoritma KNN yaitu nilai akurasi terbaik dengan K=3, K Folds = 10 memiliki akurasi yaitu 72.46% dan Algoritma KNN-PSO akurasi terbaik dengan K=3, K Folds = 10 memiliki akurasi yaitu 89.86%

#### V. Terima Kasih

Kepada seluruh rekan-rekan peneliti di Universitas Ichsan Gorontalo, terima kasih atas bimbingannya khusus rekan-rekan di Lembaga Penelitian (LEMLIT) Universitas dan kepada dosen dan peneliti di fakultas ilmu komputer, berharap peneliti dapat fokus untuk meningkatkan kemampuan lagi

#### Daftar Pustaka

- [1] Banjarsari, M. A., Budiman, I., & Farmadi, A. (2016). Penerapan K-Optimal Pada Algoritma Knn Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan Ip Sampai Dengan Semester 4. *Klik - Kumpulan Jurnal Ilmu Komputer*, 2(2), 159–173. <https://doi.org/10.20527/KLIK.V2I2.26>
- [2] Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment Analysis of Review Datasets Using Naïve Bayes' and K-NN Classifier. *International Journal of Information Engineering and Electronic Business*. <https://doi.org/10.5815/ijieeb.2016.04.07>
- [3] S. B. Imandoust and M. Bolandrafar, "Application of K-Nearest Neighbor ( KNN ) Approach for Predicting Economic Events : Theoretical Background," vol. 3, no. 5, pp. 605–610, 2013.

- [4] M. E. Lasulika, "PREDIKSI HARGA KOMODITI JAGUNG MENGGUNAKAN K-NN DAN PARTICLE SWARM OPTIMAZATION," vol. 9, pp. 233–238, 2017.
- [5] Pandie S.Y. Emerensye, Implementasi Algoritma Data Mining K-Nearest Neighbor (K-NN) Dalam Pengambilan Keputusan Pengajuan Kredit, Seminar Nasional Sains dan Teknik, Kupang. 2012
- [6] He, H., & Alippi, C. (2015). Conference report for 2014 iee symposium series on computational intelligence (IEEE SSCI 2014) [Conference Reports]. *IEEE Computational Intelligence Magazine*, 10(2), 10–12. <https://doi.org/10.1109/MCI.2015.2405273>
- [7] R. N. Whidhiasi, N. A. Wahanani and Supriyanto, Klasifikasi buah belimbing berdasarkan citra red-greenblue menggunakan knn dan Ida, Jurnal Penelitian Ilmu Komputer, *System Embedded & Logic 1(1)* : 29-35, 2013
- [8] Imandoust Bafandeh S, dkk, *Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events : Theoretical Background*, S.B Imandoust et al. *Int. Journal of Engineering Research and Applications*. Vol. 3, Issue 5, 2013
- [9] Kurniawan, D., Wibowo, W., & Astuti, P. Y. (2007a). Pemanfaatan Educational Data Mining (Edm) Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Ti-S1 Udinus). *Momentum*, 12, 48–52.
- [10] Kurniawan, D., Wibowo, W., & Astuti, P. Y. (2007b). Pemanfaatan Educational Data Mining (Edm) Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Ti-S1 Udinus). *Momentum*.
- [11] Marthasari, G. I. (2017). Implementasi Teknik Data Mining untuk Evaluasi Kinerja Mahasiswa Berdasarkan Data Akademik. *Fountain of Informatics Journal*, 2(2), 20. <https://doi.org/10.21111/fij.v2i2.1216>
- [12] Mustafa, M. S., & Simpen, I. W. (2015). Perancangan Aplikasi Prediksi Kelulusan Tepat Waktu Bagi Mahasiswa Baru Dengan Teknik Data Mining (Studi Kasus: Data Akademik Mahasiswa STMIK Dipanegara Makassar). *Creative Information Technology Journal*, 1(4), 270. <https://doi.org/10.24076/citec.2014v1i4.27>
- [13] Rohman, A. (2015). Model Algoritma K-Nearest Neighbor (K-Nn) Untuk Prediksi Kelulusan Mahasiswa. *Neo Teknika*, 1(1). <https://doi.org/10.1017/CBO9781107415324.004>
- [14] Rustam, S., Santoso, H. A., & Supriyanto, C. (2018). Optimasi K-Means Clustering Untuk Identifikasi Daerah Endemik Penyakit Menular Dengan Algoritma Particle Swarm Optimization Di Kota Semarang. *ILKOM Jurnal Ilmiah*, 10(3), 251. <https://doi.org/10.33096/ilkom.v10i3.342.251-259>
- [15] J. Han, *Data Mining: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems)*. 2011.
- [16] G. I. Marthasari, "Implementasi Teknik Data Mining untuk Evaluasi Kinerja Mahasiswa Berdasarkan Data Akademik," *Fountain Informatics J.*, vol. 2, no. 2, p. 20, 2017.

# OPTIMASI DATA MINING AKADEMIK MENGGUNAKAN ALGORITMA PARTICLE SWARM OPTIMIZATION DAN KNN UNTUK KLASIFIKASI TOPIK TUGAS AKHIR MAHASISWA

ORIGINALITY REPORT

22%

SIMILARITY INDEX

## PRIMARY SOURCES

|   |                                                                                                                                                                              |                 |
|---|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| 1 | <a href="http://jurnal.fikom.umi.ac.id">jurnal.fikom.umi.ac.id</a><br>Internet                                                                                               | 321 words — 12% |
| 2 | <a href="http://ejournal.nusamandiri.ac.id">ejournal.nusamandiri.ac.id</a><br>Internet                                                                                       | 84 words — 3%   |
| 3 | <a href="http://digilib.unmuhjember.ac.id">digilib.unmuhjember.ac.id</a><br>Internet                                                                                         | 76 words — 3%   |
| 4 | <a href="http://journal.unnes.ac.id">journal.unnes.ac.id</a><br>Internet                                                                                                     | 40 words — 1%   |
| 5 | <a href="http://researchr.org">researchr.org</a><br>Internet                                                                                                                 | 21 words — 1%   |
| 6 | <a href="http://eprints.umm.ac.id">eprints.umm.ac.id</a><br>Internet                                                                                                         | 18 words — 1%   |
| 7 | Rahmat Karim Haba, Kartika Chandra Pelangi.<br>"Pengelompokan Buah Jeruk menggunakan Naïve Bayes dan Gray Level Co-occurrence Matrix", ILKOM Jurnal Ilmiah, 2020<br>Crossref | 18 words — 1%   |
| 8 | <a href="http://www.mikroskil.ac.id">www.mikroskil.ac.id</a><br>Internet                                                                                                     | 9 words — < 1%  |
| 9 | <a href="http://search.crossref.org">search.crossref.org</a><br>Internet                                                                                                     | 8 words — < 1%  |

10

Suhardi Rustam, Haditsah Annur. "AKADEMIK DATA MINING (ADM) K-MEANS DAN K-MEANS K-NN UNTUK MENGELOMPOKAN KELAS MATA KULIAH KOSENTRASI MAHASISWA SEMESTER AKHIR", ILKOM Jurnal Ilmiah, 2019

[Crossref](#)

6 words — < 1%

EXCLUDE QUOTES ON  
EXCLUDE BIBLIOGRAPHY ON

EXCLUDE MATCHES OFF