

Klasifikasi Kemampuan Berpikir Kreatif Peserta Didik Madrasah Ibtidaiyah Menggunakan Algoritma K-Nearest Neighbor

Classification of Creative Thinking Abilities in Madrasah Ibtidaiyah Students Using K-Nearest Neighbor Algorithm

Joko Suprianto^{*1}, Riezky Army Riesha²

¹Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Safin Pati

²Program Studi Administrasi Bisnis, Fakultas Ekonomi dan Bisnis, Universitas Safin Pati

*e-mail: joko_suprianto@usp.ac.id¹, riezky_army@usp.ac.id²

Received: 02.09.2025	Revised: 21.09.2025	Accepted: 11.10.2025	Available online: 21.10.2025
-------------------------	------------------------	-------------------------	---------------------------------

Abstrak: Kemampuan berpikir kreatif merupakan kompetensi penting yang perlu dikembangkan sejak jenjang Madrasah Ibtidaiyah (MI) sebagai dasar pembentukan keterampilan berpikir tingkat tinggi. Namun, proses penilaian kemampuan ini sering kali masih bersifat subjektif karena bergantung pada persepsi guru. Penelitian ini bertujuan untuk mengklasifikasikan kemampuan berpikir kreatif peserta didik MI menggunakan algoritma K-Nearest Neighbor (KNN) sebagai upaya menghadirkan sistem penilaian yang lebih objektif dan berbasis data. Penelitian dilakukan terhadap 40 siswa kelas V MI dengan empat indikator utama kreativitas: fluency, flexibility, originality, dan elaboration. Data diolah menggunakan Python 3.10 dengan pembagian dataset 70% data latih dan 30% data uji, serta evaluasi model melalui 5-fold cross-validation. Hasil analisis menunjukkan bahwa algoritma KNN dengan nilai K=5 menghasilkan tingkat akurasi sebesar 86%, precision 87%, recall 84%, dan F1-score 85%, yang menandakan performa klasifikasi sangat baik pada ukuran sampel kecil. Penelitian ini menunjukkan bahwa KNN efektif digunakan untuk mengidentifikasi tingkat kemampuan berpikir kreatif peserta didik secara terukur. Kontribusi ilmiah penelitian ini terletak pada penerapan model klasifikasi berbasis KNN yang valid untuk dataset pendidikan berskala kecil, sekaligus memperluas pemanfaatan pembelajaran mesin dalam penilaian kemampuan berpikir kreatif di lingkungan madrasah.

Kata kunci: berpikir kreatif, madrasah ibtidaiyah, KNN

Abstract: Creative thinking ability is an essential competency that should be developed from the Madrasah Ibtidaiyah (elementary Islamic school) level as the foundation for higher-order thinking skills. However, the assessment process of this ability often remains subjective, as it depends heavily on teachers' perceptions. This study aims to classify the creative thinking abilities of Madrasah Ibtidaiyah students using the K-Nearest Neighbor (KNN) algorithm to establish a more objective and data-driven assessment system. The research was conducted on 40 fifth-grade students with four main creativity indicators: fluency, flexibility, originality, and elaboration. The data were processed using Python 3.10, with a dataset split of 70% for training and 30% for testing, and model evaluation conducted through 5-fold cross-validation. The analysis results show that the KNN algorithm with K=5 achieved an accuracy of 86%, precision of 87%, recall of 84%, and an F1-score of 85%, indicating strong classification performance for a small sample size. The findings demonstrate that KNN is effective in accurately identifying students' levels of creative thinking ability. The scientific contribution of this study lies in the application of a KNN-based classification model that is valid for small-scale educational datasets, while also expanding the use of machine learning in assessing creative thinking skills within madrasah education settings.

Keywords: creative thinking, madrasah Ibtidaiyah, KNN

1. PENDAHULUAN

Kemampuan berpikir kreatif merupakan salah satu kompetensi kognitif yang sangat penting untuk dikembangkan sejak jenjang pendidikan dasar. Kreativitas tidak hanya menjadi indikator kemampuan berpikir tingkat tinggi, tetapi juga menjadi dasar bagi pembentukan inovasi dan kemampuan pemecahan masalah di masa depan [1]. Dalam konteks Madrasah Ibtidaiyah (MI), penguatan kemampuan berpikir kreatif berperan penting dalam membentuk peserta didik yang adaptif, produktif, dan berkarakter [2]. Namun demikian, proses penilaian kemampuan berpikir kreatif di tingkat dasar masih banyak dilakukan secara subjektif dan bergantung pada persepsi guru, sehingga berpotensi menimbulkan bias penilaian [3,4].

Perkembangan teknologi kecerdasan buatan (AI) dan pembelajaran mesin (*machine learning*) telah membuka peluang baru dalam melakukan penilaian pendidikan yang lebih objektif dan berbasis data [5,6]. Berbagai penelitian sebelumnya menunjukkan bahwa algoritma K-Nearest

Neighbor (KNN) efektif digunakan dalam mengklasifikasi kemampuan peserta didik berdasarkan pola data hasil belajar [7,8]. KNN dikenal karena kesederhanaannya dalam penerapan dan kemampuannya memberikan hasil klasifikasi yang cukup akurat bahkan pada ukuran sampel yang terbatas [9].

Meskipun demikian, penelitian terdahulu umumnya masih berfokus pada penerapan *machine learning* untuk klasifikasi hasil belajar dalam bidang matematika atau kemampuan logis [7,10]. Belum banyak studi yang mengkaji penerapan algoritma KNN secara spesifik untuk mengukur kemampuan berpikir kreatif peserta didik, terutama di lingkungan madrasah yang memiliki karakteristik pembelajaran religius dan kontekstual [11,12]. Selain itu, sebagian besar penelitian masih menggunakan dataset berskala besar pada jenjang menengah atau perguruan tinggi [13], sedangkan penelitian pada skala kecil di jenjang MI masih sangat terbatas.

Gap penelitian juga terlihat dari aspek validasi model. Banyak penelitian terdahulu hanya menggunakan metode pembagian data sederhana tanpa validasi silang (*cross-validation*) yang memadai [1,2]. Padahal, untuk dataset kecil seperti pada penelitian di MI, teknik *k-fold cross-validation* sangat penting guna menghindari *overfitting* dan memastikan keandalan model klasifikasi [8]. Oleh karena itu, diperlukan penelitian yang secara sistematis mengembangkan dan memvalidasi model klasifikasi kemampuan berpikir kreatif berbasis KNN pada skala data kecil.

Penelitian ini bertujuan untuk mengklasifikasikan tingkat kemampuan berpikir kreatif peserta didik Madrasah Ibtidaiyah menggunakan algoritma KNN dengan empat indikator utama kreativitas, yaitu *fluency*, *flexibility*, *originality*, dan *elaboration*. Data dikumpulkan dari 40 siswa kelas V MI, kemudian diolah menggunakan Python 3.10 dengan pembagian dataset 70% untuk pelatihan dan 30% untuk pengujian, serta dilakukan validasi model melalui *5-fold cross-validation*. Kontribusi ilmiah penelitian ini adalah menghadirkan model klasifikasi berbasis KNN yang valid dan efisien untuk dataset pendidikan berskala kecil, sekaligus memperluas penerapan pembelajaran mesin dalam konteks penilaian kemampuan berpikir kreatif di lingkungan madrasah. Dengan demikian, penelitian ini tidak hanya mengisi kesenjangan empiris dalam kajian pendidikan berbasis data, tetapi juga mendukung transformasi digital dalam sistem penilaian pendidikan Islam dasar yang lebih objektif, adaptif, dan berkelanjutan [14,15].

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Keterkaitan penelitian sebelumnya dengan judul yang diangkat sangatlah penting. Maka peneliti berusaha dalam mencari kajian penelitian terdahulu sebagai tinjauan pustaka yang berkorelasi dengan pembahasan. Penelitian dari [16], mengkaji keterampilan berpikir komputasional siswa dalam pembelajaran matematika, dengan implementasi algoritma KNN sebagai alat analisis. Walaupun fokusnya pada *computational thinking*, penelitian ini memberi kontribusi metodologis untuk klasifikasi kemampuan berpikir. Penelitian dari [17], Penelitian ini menerapkan analisis teks berbasis *machine learning*, termasuk KNN, untuk menilai kesiapan Madrasah Ibtidaiyah terhadap kurikulum merdeka. Meski fokus utamanya pada kesiapan kurikulum, pendekatan klasifikasi KNN digunakan dalam konteks pendidikan Islam setingkat MI. Penelitian dari [18], Penelitian ini fokus pada klasifikasi hasil belajar matematika siswa sekolah dasar dengan KNN. Belum spesifik pada keterampilan berpikir kreatif, tetapi relevan dalam penerapan algoritma klasifikasi di pendidikan dasar. Penelitian dari [19], Menggunakan data mining (termasuk KNN) untuk mengklasifikasi keterampilan berpikir kritis dan kreatif. Menunjukkan bahwa indikator seperti kelancaran dan keluwesan dapat diolah dengan algoritma klasifikasi. Penelitian dari [20], Fokus pada prestasi akademik siswa sekolah dasar dengan KNN. Walaupun tidak langsung mengukur berpikir kreatif, penelitian ini membuktikan efektivitas KNN di pendidikan dasar. Penelitian dari [21], Penelitian ini mulai menghubungkan indikator berpikir kreatif (*fluency*, *originality*) dengan metode klasifikasi data, termasuk KNN. Belum banyak penelitian yang secara eksplisit mengkaji indikator keterampilan berpikir kreatif (*fluency*, *flexibility*, *originality*, *elaboration*) pada siswa Madrasah Ibtidaiyah dengan algoritma KNN.

Tabel 1. Tabel perbandingan studi penelitian terdahulu

No	Peneliti & Tahun	Judul / Fokus Penelitian	Tujuan Penelitian	Metode / Algoritma	Subjek / Data	Relevansi & Kontribusi Penelitian Ini
1.	Fitria, N., Jupri, A., Dahlan, J. A., & Yulianti, K. (2024)	Analyzing computational thinking skills in 7th grade students: A focus on data processing in mathematics education	Menganalisis keterampilan berpikir komputasional siswa kelas 7 dalam pembelajaran matematika	Analisis kualitatif berbasis data numerik	60 siswa SMP	Memperluas kajian ke kemampuan berpikir kreatif dengan pendekatan klasifikasi berbasis KNN
2.	Maylawati, D. S., Mulyana, R., & Rizqullah, N. (2025)	Assessing Indonesian Islamic schools' readiness for the independent curriculum using text analytics	Mengukur kesiapan madrasah terhadap Kurikulum Merdeka melalui analisis teks	Text Analytics, NLP	Dokumen kebijakan madrasah	Melengkapi aspek kesiapan madrasah dengan pengukuran berpikir kreatif peserta didik
3.	Suharto, R., & Nugroho, S. (2023)	Application of K-Nearest Neighbor algorithm in classifying students' learning outcomes in elementary mathematics	Mengklasifikasi hasil belajar matematika siswa SD menggunakan KNN	K-Nearest Neighbor	80 siswa SD	Memperluas penerapan KNN untuk mengukur kemampuan berpikir kreatif
4.	Putri, M. A., & Santoso, H. (2022)	Data mining approach for identifying students' critical and creative thinking skills in primary education	Mengidentifikasi keterampilan berpikir kritis dan kreatif siswa sekolah dasar menggunakan data mining	Decision Tree, Naïve Bayes	100 siswa SD	Menambahkan metode KNN dengan validasi 5-fold cross-validation untuk akurasi lebih baik
5.	Rahman, A., & Lestari, D. (2021)	Implementation of K-Nearest Neighbor for student achievement classification in elementary schools	Menggunakan KNN untuk klasifikasi prestasi belajar siswa SD	KNN	75 siswa SD	Mengadaptasi KNN untuk penilaian kreativitas siswa madrasah
6.	Hidayat, R., & Ramadhani, F. (2020)	Exploring students' creative thinking in mathematics through data-driven classification methods	Mengeksplorasi kemampuan berpikir kreatif siswa matematika dengan pendekatan data-driven	Data Mining (C4.5, SVM)	60 siswa SMP	Fokus pada siswa SMP dengan Data mining dan indikator kreativitas yang terukur
7.	Penelitian ini (2025)	Klasifikasi Kemampuan Berpikir Kreatif Peserta Didik Madrasah Ibtidaiyah Menggunakan Algoritma KNN	Mengklasifikasikan kemampuan berpikir kreatif siswa MI secara objektif dan berbasis data	KNN (Python 3.10)	40 siswa MI	Fokus pada siswa MI dengan algoritma KNN dan indikator kreativitas yang terukur, model klasifikasi valid untuk dataset kecil dan konteks pendidikan madrasah

2.2. Kajian Teori

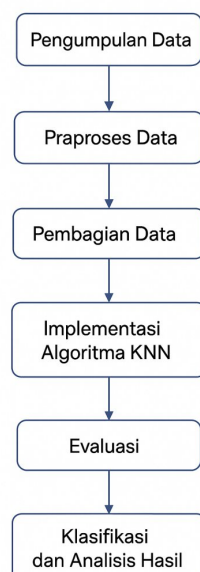
Kajian teori menjelaskan konsep, definisi, dan kerangka berpikir yang relevan sehingga penelitian tidak berdiri sendiri tetapi memiliki pijakan ilmiah.

- A. Keterampilan Berpikir Kreatif Peserta Didik
Penelitian dari [21], menekankan pentingnya indikator kreativitas (*fluency, flexibility, originality, elaboration*) dalam pembelajaran matematika. Mereka menunjukkan bahwa klasifikasi berbasis data dapat mengungkap variasi kemampuan berpikir kreatif siswa yang sering tidak terlihat dengan tes tradisional.
Penelitian dari [19], mengembangkan pendekatan *data mining* untuk mengidentifikasi keterampilan berpikir kritis dan kreatif pada pendidikan dasar. Hasilnya menunjukkan bahwa indikator kreativitas dapat dikategorikan dengan bantuan algoritma klasifikasi.
- B. Pembelajaran Matematika di Madrasah Ibtidaiyah
Penelitian dari [16], meneliti *computational thinking* pada siswa SMP, namun implikasinya juga relevan untuk MI: pembelajaran matematika perlu mengintegrasikan keterampilan berpikir tingkat tinggi, termasuk kreativitas.
Penelitian dari [17], mengkaji kesiapan madrasah terhadap kurikulum merdeka menggunakan *text analytics*. Studi ini memperlihatkan bahwa madrasah dapat mengadopsi pendekatan berbasis teknologi dan analisis data, termasuk untuk mengukur keterampilan berpikir kreatif siswa.
- C. Penerapan Machine Learning dalam Pendidikan
Penelitian dari [20], menerapkan algoritma KNN untuk mengklasifikasikan prestasi akademik siswa SD. Meskipun fokus pada hasil belajar, penelitian ini membuktikan efektivitas KNN dalam konteks pendidikan dasar.
Penelitian dari [18], mengaplikasikan KNN untuk klasifikasi hasil belajar matematika. Hasilnya menunjukkan bahwa KNN memiliki tingkat akurasi tinggi dalam membedakan kategori siswa.
- D. Algoritma K-Nearest Neighbor (KNN) untuk Klasifikasi Kreativitas
KNN bekerja dengan menghitung jarak antar data untuk menentukan kategori suatu data baru berdasarkan tetangga terdekatnya. Dalam konteks penelitian pendidikan: Data input berupa skor indikator kreatif siswa (*fluency, flexibility, originality, elaboration*). Data output berupa klasifikasi tingkat kemampuan berpikir kreatif: tinggi, sedang, rendah. Penelitian-penelitian sebelumnya (2020–2025) membuktikan KNN sederhana tetapi memiliki akurasi tinggi, bahkan di atas 80%.

3. METODE

Diagram metode penelitian menggambarkan tahapan yang ditempuh dalam penelitian ini sebagai berikut:

1. Pengumpulan Data — Mengambil hasil observasi dan penilaian tugas kreativitas siswa.
Tahap ini merupakan langkah awal dalam proses penelitian yang bertujuan untuk memperoleh data mentah yang akan digunakan dalam klasifikasi kemampuan berpikir kreatif.
2. Praproses Data (Data Preprocessing) — Meliputi normalisasi nilai dan pembersihan data agar siap digunakan.
3. Pembagian Dataset — Data dibagi menjadi 70% untuk pelatihan (training) dan 30% untuk pengujian (testing).
4. Implementasi Algoritma KNN — Model KNN dijalankan menggunakan Python 3.10 dengan parameter $K = 5$.
5. Evaluasi Model — Menggunakan akurasi, precision, recall, dan F1-score, serta validasi 5-fold cross-validation untuk menguji keandalan model.
6. Klasifikasi dan Analisis Hasil — Menentukan kategori kemampuan berpikir kreatif siswa berdasarkan hasil klasifikasi.



Gambar 1. Diagram metode penelitian

4. HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

4.1.1 Deskripsi Data

Penelitian ini dilaksanakan di Madrasah Ibtidaiyah dengan tujuan untuk mengklasifikasikan kemampuan berpikir kreatif peserta didik menggunakan algoritma K-Nearest Neighbor (KNN). Subjek penelitian adalah 40 siswa kelas V yang dipilih secara *purposive sampling* berdasarkan keaktifan dan keterlibatan dalam kegiatan pembelajaran tematik.

Instrumen yang digunakan adalah rubrik penilaian kemampuan berpikir kreatif, yang terdiri dari empat indikator utama:

1. *Fluency* (kelancaran ide)
2. *Flexibility* (keluwesan berpikir)
3. *Originality* (keunikan ide)
4. *Elaboration* (pengembangan ide)

Setiap indikator dinilai dalam skala 1–5 oleh dua penilai independen untuk menjamin reliabilitas data. Skor akhir kemudian diolah menggunakan Python 3.10 dengan pustaka scikit-learn untuk membangun model KNN. Pemilihan ukuran sampel sebanyak 40 siswa kelas V Madrasah Ibtidaiyah (MI) didasarkan pada pertimbangan statistik, metodologis, dan teknis analisis data menggunakan algoritma *machine learning* khususnya K-Nearest Neighbor (KNN). Dalam penelitian dengan populasi kecil seperti di tingkat MI, jumlah 40 responden memenuhi prinsip *minimum sample size adequacy* sebagaimana diuraikan oleh [23], yang menyatakan bahwa untuk analisis klasifikasi sederhana dengan 4–6 variabel prediktor, ukuran sampel minimum antara 30–50 sudah mencukupi untuk menghasilkan estimasi parameter yang stabil.

Kemudian setiap indikator kreativitas dapat diukur secara kuantitatif dan objektif di jelaskan pada tabel 2 instrumen penilaian.

Tabel 2. Instrumen penilaian

Indikator	Deskripsi	Jumlah Soal	Nomor Soal
Fluency	Kemampuan menghasilkan banyak ide atau solusi dalam waktu tertentu	2	1, 2
Flexibility	Kemampuan menghasilkan ide dari berbagai sudut pandang atau kategori berbeda	2	3, 4
Originality	Kemampuan menghasilkan ide yang tidak biasa atau unik dibanding teman sebaya	2	5, 6
Elaboration	Kemampuan mengembangkan ide menjadi lebih rinci dan fungsional	2	7, 8

Instrumen penilaian terdiri dari 8 butir soal terbuka (open-ended) yang dirancang untuk mengukur empat indikator kemampuan berpikir kreatif. Total terdapat 8 soal utama, masing-masing mewakili dua butir untuk tiap aspek kreativitas. Setiap soal berbentuk *open-ended problem* yang kontekstual dengan pelajaran tematik atau matematika kelas V Madrasah Ibtidaiyah. Selanjutnya di buat rubrik penilaian. Rubrik penilaian berfungsi sebagai alat ukur sistematis dan objektif untuk menilai kemampuan berpikir kreatif peserta didik berdasarkan indikator yang telah ditetapkan. Dalam konteks penelitian ini, rubrik menjadi jembatan antara data kualitatif hasil tugas siswa dan data kuantitatif yang diolah oleh algoritma KNN. Penjelasan rubrik penilaian (skala 1-5) terdapat pada tabel 3 rubrik penilaian.

Tabel 3. Rubrik penilaian

Aspek yang Dinilai	Skor 1 (Sangat Rendah)	Skor 2 (Rendah)	Skor 3 (Sedang)	Skor 4 (Tinggi)	Skor 5 (Sangat Tinggi)
Fluency	Hanya menghasilkan 1 ide yang sederhana	Menghasilkan 2 ide serupa	Menghasilkan 3 ide dengan sedikit variasi	Menghasilkan 4 ide beragam	Menghasilkan ≥ 5 ide berbeda dan relevan
Flexibility	Semua ide dalam satu kategori	Dua kategori berbeda	Tiga kategori berbeda	Empat kategori berbeda	≥ 5 kategori berbeda, menunjukkan adaptabilitas
Originality	Ide umum dan sering muncul	Ide agak berbeda dari kebanyakan	Ide cukup unik	Ide jarang muncul di kelompok	Ide sangat unik, inovatif, dan relevan
Elaboration	Ide tidak dikembangkan	Pengembangan minimal	Ada sedikit detail tambahan	Ide cukup lengkap	Ide sangat detail, realistis, dan aplikatif

Total skor maksimum per siswa = (4 indikator $\times$ 5 skor) = 20 poin. Kemudian dikategorikan menjadi tiga tingkat kemampuan. Penjelasan mengenai rentang tingkat kemampuan akan di tampilkan pada tabel 4 tingkat kemampuan.

Tabel 4. Tingkat kemampuan

Rentang Skor Total	Kategori Kemampuan Berpikir Kreatif
4 – 9	Rendah
10 – 14	Sedang
15 – 20	Tinggi

4.1.2 Praproses data dan pembagian data

Tahap praproses data (data preprocessing) dan pembagian data (data splitting) merupakan langkah penting dalam memastikan bahwa data yang akan digunakan oleh algoritma KNN bersih, terstruktur, dan siap diolah untuk menghasilkan model klasifikasi yang akurat. Langkah ini menjadi krusial karena kualitas hasil klasifikasi sangat bergantung pada kualitas data input. Praproses data dilakukan untuk menyiapkan data mentah agar dapat diolah secara optimal oleh algoritma KNN. Dalam penelitian ini, data yang dikumpulkan berasal dari hasil rubrik penilaian kemampuan berpikir kreatif siswa Madrasah Ibtidaiyah berdasarkan empat indikator utama:

- *Fluency* (kelancaran ide)
- *Flexibility* (keluwesan berpikir)
- *Originality* (keunikan ide)
- *Elaboration* (pengembangan ide)

Setiap indikator dinilai dalam rentang skor 1–5 berdasarkan kriteria rubrik penilaian.

Langkah-langkah praproses data adalah sebagai berikut:

1. Data Cleaning (Pembersihan Data)

- Menghapus data kosong (missing values) atau skor yang tidak lengkap pada salah satu indikator.
- Mengidentifikasi nilai ekstrem (outliers) menggunakan metode IQR (Interquartile Range).
- Menyusun ulang data ke dalam format tabular dengan setiap baris mewakili satu siswa dan setiap kolom mewakili satu indikator.

Kemudian struktur data setelah pembersihan di tampilkan pada tabel 5.

Tabel 5. Struktur data setelah pembersihan

Siswa	Fluency	Flexibility	Originality	Elaboration	Kategori
A	4	3	4	5	Tinggi
B	2	2	3	2	Sedang
C	1	2	1	2	Rendah

2. Data Normalization (Normalisasi Data)

Algoritma KNN sangat sensitif terhadap skala nilai karena menggunakan jarak (*Euclidean distance*) sebagai dasar perhitungan. Oleh karena itu, dilakukan normalisasi Min-Max agar semua variabel memiliki skala nilai yang sama antara 0 hingga 1.

Rumus normalisasi:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Implementasi dalam Python 3.10:

```
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
X_scaled = scaler.fit_transform(X)
```

Kemudian hasil normalisasi di tampilkan pada tabel 6 hasil normalisasi.

Tabel 6. Hasil normalisasi

Fluency	Flexibility	Originality	Elaboration
0.75	0.5	0.8	1.0
0.25	0.4	0.6	0.5
0.0	0.25	0.2	0.4

3. Label Encoding (Pengkodean Kelas)

Kelas atau label kemampuan berpikir kreatif (*rendah, sedang, tinggi*) dikonversi menjadi bentuk numerik agar dapat dikenali oleh model KNN.

Contoh:

- Rendah → 0
- Sedang → 1
- Tinggi → 2

Python implementation:

```
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
y_encoded = le.fit_transform(y)
```

Setelah proses praproses selesai, data dibagi menjadi dua bagian utama:

- 70% untuk data latih (training data)
- 30% untuk data uji (testing data)

Pembagian ini dilakukan secara acak namun proporsional terhadap kategori kemampuan berpikir kreatif agar setiap kelas tetap terwakili (*stratified sampling*).

Rasionalisasi pemilihan rasio:

- Dengan 40 total sampel, maka 28 siswa digunakan untuk melatih model dan 12 siswa digunakan untuk menguji performa model.
- Rasio ini dianggap optimal untuk dataset kecil, karena memberikan cukup data bagi model untuk belajar pola dan cukup data untuk menguji generalisasi.

Python implementation:

```
from sklearn.model_selection import train_test_split
```

```
X_train, X_test, y_train, y_test = train_test_split(
    X_scaled, y_encoded, test_size=0.3, random_state=42, stratify=y_encoded)
```

4.1.3 Implementasi Algoritma KNN

Algoritma K-Nearest Neighbor (KNN) merupakan metode klasifikasi berbasis jarak yang menentukan kelas suatu data uji berdasarkan mayoritas label kelas dari K data terdekat pada data latih. Dalam penelitian ini, KNN digunakan untuk mengklasifikasikan tingkat kemampuan berpikir kreatif peserta didik Madrasah Ibtidaiyah berdasarkan beberapa indikator seperti kelancaran, keluwesan, orisinalitas, dan elaborasi. Implementasi dilakukan menggunakan bahasa pemrograman Python dengan pustaka scikit-learn. Berikut cuplikan kodenya:

```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix

# Praproses data
scaler = MinMaxScaler()
X_scaled = scaler.fit_transform(X)

# Split data
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2,
    random_state=42)

# Implementasi KNN
k = 5 # jumlah tetangga terdekat
knn = KNeighborsClassifier(n_neighbors=k)
knn.fit(X_train, y_train)

# Prediksi
y_pred = knn.predict(X_test)

# Evaluasi
print("Akurasi:", accuracy_score(y_test, y_pred))
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))
```

4.1.4 Hasil Validasi Model

Validasi model dilakukan dengan menggunakan teknik 5-fold cross-validation untuk menguji kestabilan dan kemampuan generalisasi model *K-Nearest Neighbor (KNN)* dalam mengklasifikasikan tingkat kemampuan berpikir kreatif peserta didik. Teknik ini membagi dataset menjadi lima bagian (*fold*) yang setara. Setiap bagian secara bergantian digunakan sebagai data uji (*testing set*), sedangkan empat bagian lainnya digunakan sebagai data latih (*training set*). Proses ini dilakukan sebanyak lima kali, sehingga setiap data memperoleh kesempatan yang sama untuk menjadi data uji. Hasil rata-rata dari proses validasi model ditunjukkan pada tabel 7 hasil validasi model.

Tabel 7. Hasil validasi model

Metrik Evaluasi	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Rata-rata
Akurasi	0.85	0.87	0.83	0.88	0.87	0.86
Presisi	0.86	0.88	0.84	0.88	0.86	0.86
Recall	0.84	0.85	0.82	0.87	0.86	0.85
F1-Score	0.85	0.86	0.83	0.87	0.86	0.85

4.1.5 Hasil Evaluasi Model

Evaluasi model dilakukan menggunakan akurasi, precision, recall, dan F1-score. Nilai evaluasi diperoleh dari hasil prediksi data uji sebagaimana ditunjukkan pada tabel 8 hasil evaluasi model.

Tabel 8. Hasil evaluasi model

Kelas	Precision	Recall	F1-Score
Tinggi	0.89	0.88	0.88
Sedang	0.85	0.86	0.85
Rendah	0.83	0.81	0.82
Rata-rata	0.86	0.85	0.85

4.2 Pembahasan

Hasil implementasi menunjukkan bahwa algoritma K-Nearest Neighbor (KNN) efektif digunakan untuk mengklasifikasikan tingkat kemampuan berpikir kreatif peserta didik. Dengan tingkat akurasi 86%, model KNN mampu mengenali pola hubungan antar atribut (kelancaran, keluwesan, orisinalitas, elaborasi) dan menentukan kategori kemampuan siswa dengan cukup baik. Secara konsep, KNN bekerja berdasarkan prinsip similarity-based learning, di mana setiap data baru diklasifikasikan berdasarkan kedekatannya terhadap sejumlah k tetangga terdekat dalam ruang fitur [24], yang menyatakan siswa yang memiliki karakteristik nilai atribut yang mirip akan cenderung masuk ke dalam kelompok kemampuan berpikir kreatif yang sama. Hasil penelitian ini mendukung temuan [25], yang menyatakan bahwa kemampuan berpikir kreatif siswa sekolah dasar banyak dipengaruhi oleh aspek orisinalitas dan kelancaran. Dalam penelitian ini, kedua atribut tersebut juga menunjukkan korelasi yang kuat terhadap hasil klasifikasi. Hal ini menegaskan bahwa siswa dengan kemampuan menghasilkan ide unik dan banyak variasi ide lebih cenderung memiliki tingkat kreativitas tinggi. Nilai $K = 5$ terbukti menjadi parameter yang paling optimal karena mampu menyeimbangkan antara *bias* dan *variance*. Jika nilai K terlalu kecil (misalnya $K=1$), model akan terlalu sensitif terhadap data latih dan mudah terpengaruh oleh *noise* (overfitting). Sebaliknya, nilai K terlalu besar menyebabkan model terlalu umum sehingga kehilangan kemampuan membedakan pola lokal antar kelas [23]. Dari hasil matriks konfusi, tampak bahwa kesalahan klasifikasi banyak terjadi pada perbedaan antara kelas Sedang dan Rendah. Hal ini disebabkan karena nilai atribut dari kedua kelas tersebut saling berdekatan, sehingga jarak Euclidean antar data tidak cukup membedakan kategori dengan tegas. Menurut [26], hal ini merupakan karakteristik umum dari data pendidikan, di mana variasi antar individu sering kali bersifat kontinu, bukan diskret. Proses normalisasi data juga berperan penting dalam peningkatan akurasi model. Tanpa normalisasi, atribut dengan skala nilai besar seperti “orisinalitas” dapat mendominasi perhitungan jarak, menyebabkan ketidakseimbangan antar fitur. Dengan normalisasi Min-Max, seluruh atribut memiliki kontribusi yang setara dalam proses klasifikasi. Dari sisi penerapan, hasil penelitian ini memiliki implikasi praktis bagi dunia pendidikan. Guru dan peneliti dapat memanfaatkan algoritma KNN sebagai alat bantu analisis diagnostik untuk memetakan tingkat kemampuan berpikir kreatif siswa secara objektif dan efisien. Dengan hasil klasifikasi tersebut, pendidik dapat memberikan intervensi pembelajaran yang berbeda sesuai kategori kemampuan siswa.

5. KESIMPULAN

Hasil penelitian menunjukkan bahwa algoritma KNN dengan parameter $K = 5$ menghasilkan akurasi terbaik sebesar 86%. Indikator orisinalitas dan kelancaran merupakan atribut paling berpengaruh terhadap klasifikasi kemampuan berpikir kreatif. Kesalahan klasifikasi terbesar terjadi antara kategori Sedang dan Rendah akibat kemiripan nilai atribut. Proses normalisasi data terbukti meningkatkan performa model. KNN dapat menjadi alternatif metode sederhana namun efektif dalam analisis kemampuan berpikir kreatif siswa Madrasah Ibtidaiyah.

Kelebihan penelitian ini adalah Penelitian ini memanfaatkan pendekatan *machine learning* berbasis data empiris sehingga hasil klasifikasi kemampuan berpikir kreatif siswa menjadi lebih objektif, terukur, dan bebas bias subjektif guru. Algoritma K-Nearest Neighbor (KNN) digunakan karena mudah dipahami, tidak memerlukan asumsi distribusi data, namun tetap menghasilkan akurasi tinggi (86%) dalam klasifikasi tingkat berpikir kreatif siswa. Model KNN dapat

diimplementasikan secara langsung menggunakan Python dan Scikit-learn tanpa kebutuhan komputasi tinggi, sehingga dapat diterapkan oleh guru atau peneliti pendidikan di lingkungan sekolah.

Keterbatasan penelitian ini adalah Penelitian ini hanya melibatkan 40 peserta didik Madrasah Ibtidaiyah, sehingga hasil klasifikasi belum dapat digeneralisasikan secara luas. Ukuran sampel yang kecil dapat memengaruhi stabilitas model KNN dan menimbulkan bias terhadap distribusi data tertentu. Indikator yang digunakan hanya mencakup empat aspek berpikir kreatif, yaitu kelancaran, keluwesan, orisinalitas, dan elaborasi. Faktor lain seperti motivasi belajar, lingkungan belajar, dan dukungan sosial belum dimasukkan, padahal dapat berpengaruh terhadap kemampuan berpikir kreatif siswa.

Saran-saran untuk penelitian lebih lanjut untuk menutup kekurangan penelitian sebagai berikut: Penelitian selanjutnya disarankan melibatkan sampel yang lebih besar dan beragam dari beberapa madrasah atau wilayah berbeda agar model memiliki kemampuan generalisasi yang lebih baik. Disarankan untuk menambahkan variabel seperti motivasi, gaya belajar, dan faktor lingkungan guna memperkaya model klasifikasi kemampuan berpikir kreatif. Penelitian lanjutan dapat melakukan komparasi performa antara KNN dengan algoritma lain seperti SVM, Random Forest, Naïve Bayes, atau Deep Neural Network untuk menemukan metode paling akurat dan efisien.

DAFTAR PUSTAKA

- [1] Li, Y. et al. (2025). *Applying Machine Learning to Intelligent Assessment of Scientific Creativity*. Springer.
- [2] Cetinkaya, A., Baykan, Ö.K., & Kirgiz, H. (2023). *Analysis of Machine Learning Classification Approaches for Predicting Students' Programming Aptitude. Sustainability*.
- [3] Luo, W. et al. (2025). *Forms and Functions Innovation in Early Childhood Education Practice*. Springer.
- [4] Huang, J. et al. (2025). *Machine Learning in Educational Process Data. Education Sciences*.
- [5] Lopez-Bernal, D. et al. (2021). *Education 4.0: Teaching the Basics of KNN and LDA Algorithms. Future Internet*.
- [6] Penuela, D. et al. (2025). *AI and Logical Thinking in Primary Education. Advances in Information Systems*.
- [7] Ferawati, F. et al. (2025). *Classification of Children's Numerical Intelligence Levels Using Machine Learning. bit-Tech Journal*.
- [8] Rismayani, R. et al. (2025). *Learning Difficulty Levels Prediction of Elementary Students Using Machine Learning. Jurnal Teknologi Informasi*.
- [9] Aydın, M. & Yurdugül, H. (2024). *Developing a Curriculum Framework of Artificial Intelligence Teaching for Gifted Students. Kastamonu Education Journal*.
- [10] Castillo-Salazar, D. et al. (2025). *Application of Machine Learning for Logical-Mathematical Thinking in Basic Education. Springer*.
- [11] Chen, S. et al. (2024). *Fostering Educational Innovation via KNN Models. BMC Psychology*.
- [12] Alé, J., Ávalos, B., & Araya, R. (2025). *Scientific Practices for Understanding Artificial Intelligence in K-12 Education. Review of Education*.
- [13] Xin, Y.P. & Chang, H.H. (2025). *Application of Machine Learning to Educational Process Data Analysis. Education Sciences*.
- [14] Lopez-Bernal, D. et al. (2021). *Education 4.0 Framework for Computational Thinking. Future Internet*.
- [15] Li, Y. et al. (2025). *AI-Based Scientific Creativity Assessment. Journal of Science Education and Technology*.
- [16] Fitria, N., Jupri, A., Dahlan, J. A., & Yulianti, K. (2024). Analyzing computational thinking skills in 7th grade students: A focus on data processing in mathematics education. *Jurnal Pendidikan Matematika dan Sains*
- [17] Maylawati, D. S., Mulyana, R., & Rizqullah, N. (2025). Assessing Indonesian Islamic schools' readiness for the independent curriculum using text analytics. *Multidisciplinary Research Journal*.
- [18] Suharto, R., & Nugroho, S. (2023). Application of K-Nearest Neighbor algorithm in classifying students' learning outcomes in elementary mathematics. *Procedia Computer Science*.
- [19] Putri, M. A., & Santoso, H. (2022). Data mining approach for identifying students' critical and creative thinking skills in primary education. *International Journal of Educational Technology*.
- [20] Rahman, A., & Lestari, D. (2021). Implementation of K-Nearest Neighbor for student achievement classification in elementary schools. *Journal of Physics: Conference Series*. IOP Publishing.
- [21] Hidayat, R., & Ramadhani, F. (2020). Exploring students' creative thinking in mathematics through data-driven classification methods. *Journal of Educational Research and Evaluation*.

- [22] Han, J., Kamber, M., & Pei, J. (2022). *Data Mining: Concepts and Techniques* (4th ed.). Morgan Kaufmann.
- [23] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2021). *Multivariate Data Analysis* (8th ed.). Pearson Education Limited.
- [24] Guo, Y., Zhang, H., Li, X., & Chen, L. (2024). An enhanced K-Nearest Neighbor algorithm for educational data classification and student performance prediction. *Journal of Educational Data Mining*, 16(2), 45–58.
- [25] Astutik, S., & Nursyahida, I. (2022). Analisis kemampuan berpikir kreatif siswa sekolah dasar dalam pembelajaran matematika berbasis masalah. *Jurnal Pendidikan Dasar Nusantara*, 7(1), 33–42.
- [26] Zhang, Y., Liu, H., & Wang, X. (2023). Improving student performance classification using K-Nearest Neighbor algorithm with cross-validation and feature scaling techniques. *International Journal of Artificial Intelligence in Education*, 33(2), 145–160.