

## BAB I PENDAHULUAN

### 1.1. Latar Belakang

Penerapan pembelajaran mesin (*machine learning*) belakangan ini masih dilakukan di berbagai lini kehidupan. Adapun pada penelitian ini juga mencoba menerapkan pembelajaran mesin untuk mengenali motif-motif batik melalui pengolahan citra (*image processing*). Hal mendasar yang melatarbelakangi penelitian ini adalah sebagai upaya mendukung generasi muda dalam melestarikan budaya kain batik khas Indonesia sebagaimana tren berbusana batik belakangan ini di kalangan anak muda walaupun masih didasari atas alasan estetika [1].

Adapun untuk memperjelas kerangka penelitian, telah dibuat beberapa batasan yang ada pada produk klasifikasi motif batik ini agar proses penelitian dapat dipahami secara rinci dan objektif. Berikut adalah batasan-batasannya.

- a. Dataset motif batik yang digunakan dalam penelitian adalah motif Parang dan motif Ceplok dengan masing-masing jumlah 40 gambar untuk pelatihan (*training*) dan 9 gambar untuk pengujian (*testing*);
- b. Proses input gambar motif yang hendak diidentifikasi perlu diinputkan terlebih dahulu oleh pengguna pada aplikasi web yang sudah dibangun dan terintegrasi dengan model-model pembelajaran;
- c. Tingkat kecerahan, ketajaman, dan jelasnya berkas gambar yang diinputkan ke dalam web diasumsikan sudah diatur sebelumnya;
- d. Ekstensi berkas gambar yang digunakan pada penelitian adalah JPG/JPEG;
- e. Ukuran dimensi berkas gambar yang digunakan adalah 360 piksel  $\times$  360 piksel.

Tujuan dari pengembangan model klasifikasi motif batik secara khusus adalah untuk mengetahui tingkat performa dari tiap-tiap model (*K-Nearest Neighbors*, *Support Vector Classifier*, *Naïve Bayes*) yang digunakan dalam menyelesaikan tugas klasifikasi motif batik dan secara luas bertujuan untuk membantu kaum muda dalam mengenali motif-motif batik yang ada di sekitar mereka melalui aplikasi web.

Ditinjau dari perspektif teknis, penelitian ini dapat digunakan sebagai referensi bagi penelitian berikutnya yang juga meneliti mengenai klasifikasi motif batik Indonesia. Dengan mengetahui hasil performa model pada penelitian ini, ke depannya bisa dilakukan optimalisasi performa melalui konfigurasi yang berbeda dan lebih presisi pada parameter model. Selain itu, diharapkan kebermanfaatannya dari penelitian ini juga dapat dirasakan oleh masyarakat terlebih generasi muda yang ingin mengenal motif batik Indonesia dengan mudah.

## 1.2. Profil Mitra Studi Independen

PT. Bisa Artifisial Indonesia merupakan perusahaan rintisan digital berbasis *education technology (edutech)* yang fokus pada pembelajaran terkait kecerdasan artifisial (AI), ilmu data (*data science*) dan subsektor lainnya. Sebagaimana visi dari perusahaan yaitu, “Menjadi startup terdepan di Indonesia di bidang pembelajaran artificial intelligence (AI)”.

Dalam program studi independen Teknologi dan Riset Kecerdasan Artifisial, secara umum berfokus pada mempersiapkan mahasiswa untuk memiliki kompetensi keahlian sebagai seorang *programmer* di bidang *data science*, *machine learning*, dan juga *artificial intelligence*. Konsentrasi turunan yang dipelajari yaitu mengenai ilmu pemrosesan citra (*image processing*), penambangan data (*data mining*), dan pembelajaran mesin (*machine learning*).

Semua kegiatan pembelajaran dilakukan secara sinkronus daring melalui platform TAMPIL.ID. Berikut adalah jadwal kegiatan pembelajaran yang dilaksanakan selama program studi independen Teknologi dan Riset Kecerdasan Artifisial (*AI-Hacker*) di PT. Bisa Artifisial Indonesia.

Tabel 1.2.1 Jadwal pembelajaran studi independen

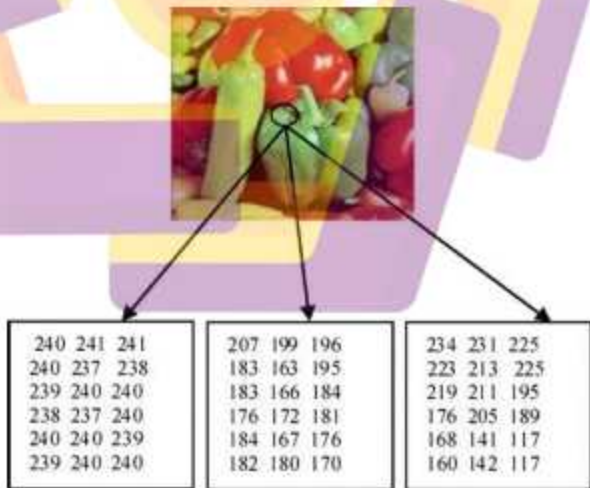
No.	Nama Kegiatan	Pelaksanaan	
		Hari	Durasi
1	Pembelajaran terjadwal – <i>image processing</i>	Senin	08.00 – 10.30 WIB
2	Pembelajaran terjadwal – <i>Data mining</i>	Senin	15.30 – 18.00 WIB
3	Pembelajaran terjadwal – <i>Machine learning</i>	Selasa	13.00 – 15.30 WIB
4	Pembelajaran terjadwal – <i>Data science</i>	Sabtu	08.00 – 10.30 WIB

5	Proyek industri – Mentoring BisaTani	Rabu	09.00 – 11.00 WIB
6	Pembelajaran tamu	Jumat-Minggu	-
7	Pembelajaran mandiri	<i>tidak ditentukan</i>	-

### 1.3. Landasan Teori

#### 1.3.1. Citra dan Pengolahan Citra Digital

Citra digital merupakan suatu gambaran atau kemiripan dari sebuah objek yang dapat diolah oleh komputer [2]. Adapun menurut Sumijan dan Pradeni Ayu Widya Purnama (2021) mendefinisikan citra digital merupakan larik (*array*) dua dimensi yang mana nilai  $f(x,y)$  telah dikonversi ke dalam bentuk diskrit baik pada koordinat citra maupun kecerahannya [3]. Kemudian, pengolahan citra digital (*image processing*) didefinisikan sebagai suatu bentuk pengolahan atau pemrosesan sinyal dengan inputan berkas gambar (*image*) dan ditransformasikan menjadi gambar lain sebagai output dengan teknik tertentu [4].



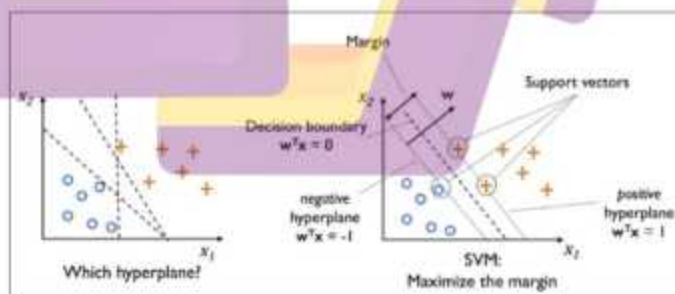
Gambar 1.3.1.1 Citra dan konsep matriks RGB

### 1.3.2. K-Nearest Neighbors

*K-Nearest Neighbors* (KNN) merupakan sebuah algoritma untuk pengklasifikasi pembelajaran terawasi (*supervised learning*) non-parametrik, dimana proses klasifikasi atau prediksi pengelompokkan suatu data didasari pada jarak kedekatan (nilai konstanta  $k$ ) [5]. Algoritma KNN mengimplementasikan pembelajaran berdasarkan tetangga terdekat dari setiap titik data, di mana nilai konstanta  $k$  yang merupakan bilangan bulat ditentukan sendiri oleh pengguna [6]. Konfigurasi nilai  $k$  yang paling optimal sangat dipengaruhi oleh dataset.

### 1.3.3. Support Vector Classifier

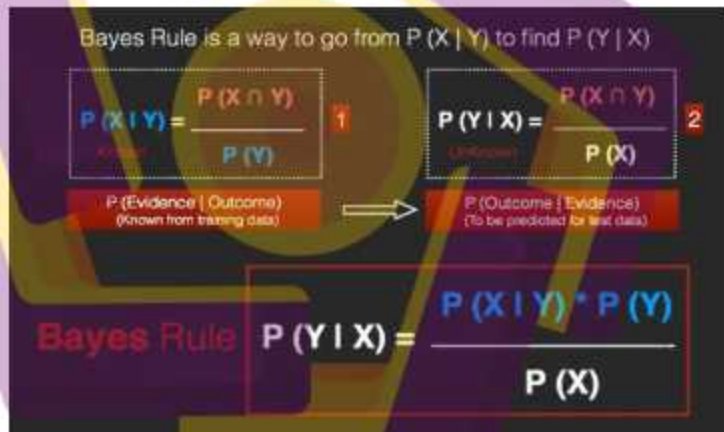
*Support Vector Classifier* (SVC) merupakan algoritma turunan dari *Support Vector Machine* (SVM) yang khusus menangani tugas-tugas klasifikasi. SVM merupakan model pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah fitur yang berdimensi tinggi dan dilatih dengan menggunakan algoritma pembelajaran yang didasarkan pada teori optimasi [7]. Algoritma SVM akan berusaha mencari *hyperplane* terbaik yang digunakan untuk memisahkan dua kelas atau lebih [8]. Pada ruang dua dimensi, fungsi klasifikasinya berbentuk garis. Sedangkan pada ruang tiga dimensi, fungsi klasifikasinya berbentuk bidang.



Gambar 1.3.3.1 Mekanisme kerja *hyperplane* pada SVM

### 1.3.4. Naïve Bayes

*Naïve Bayes* adalah sebuah algoritma untuk mengklasifikasikan suatu variabel dengan metode probabilitas dan statistika [9]. *Naïve Bayes* menerapkan cabang matematika yaitu teori probabilitas untuk mencari peluang terbesar dari kemungkinan hasil klasifikasi dengan cara melihat frekuensi tiap klasifikasi pada data pelatihan (*data training*). Karena *naïve bayes* merupakan model yang menggunakan metode probabilistik, menjadikannya dapat dikodekan dengan mudah dan pemrosesan prediksi yang sangat cepat [10]. Dengan pertimbangan kecepatan pemrosesan tersebut menjadikan *naïve bayes* algoritma pilihan untuk pengembangan aplikasi yang diperlukan untuk merespons permintaan pengguna secara instan [10].



Gambar 1.3.4.1 Rumus probabilitas dalam teori bayes

### 1.3.5. Perangkat Lunak Pendukung

Dalam proses pengerjaan penelitian ini terdapat berbagai aplikasi perangkat lunak yang digunakan sebagai penunjang proyek. Untuk proses penyusunan kode program, penelitian ini menggunakan perangkat lunak *Visual Studio Code* versi 1.73 dengan didukung beberapa ekstensi.

### 1.3.6. Bahasa Pemrograman

Bahasa pemrograman yang digunakan dalam penelitian ini adalah Python versi 3.9.7, baik dalam proses pembangunan model pembelajaran maupun dalam pembangunan platform *deployment* (web). Aplikasi web yang digunakan sebagai layar interaksi dengan pengguna, dibangun melalui bantuan *framework* pengembangan web Python yaitu Flask.

### 1.3.7. Dokumentasi

Untuk mendokumentasikan semua tahapan yang dilakukan selama penelitian termasuk di dalamnya penambahan, perubahan, ataupun penghapusan baris kode program maka digunakan *version control system* (VCS) Github. Repositori penelitian ini dapat diakses secara publik pada pranala berikut <https://github.com/katibposha/batik-classify-flask>.

