

BAB I

PENDAHULUAN

1.1 Latar Belakang

Negara agraris menurut (Hasugian et al., 2021) merupakan sebutan untuk negara Indonesia, sebab mayoritas penduduknya menopangkan kehidupan mereka sebagai petani. Dalam pemenuhan pasokan bahan pangan bidang pertanian masih berperan sangat penting, karena beras merupakan komoditas penting dan rata-rata penduduk Indonesia memilih beras sebagai makanan pokok dalam pemenuhan kebutuhan pangan sehari-hari dengan tingkat konsumsi sebesar 132,98 kg/kapita/tahun (Hasugian et al., 2021). Berdasarkan data yang dihimpun oleh *United States Departement Agriculture (USDA)*, diperkirakan produksi beras dunia mencapai 503,27 juta metrik ton (MT) pada musim panen 2022/2023, yang mengalami penurunan dari musim 2021/2022 sekitar 11,78 juta MT atau sebesar (2,29%) (Kusnandar, 2023). Sebagai kawasan produsen beras terbesar keempat dunia dan nomor satu di Asia Tenggara, Indonesia memproduksi sekitar 34,6 juta metrik ton (MT) pada musim 2022/2023 yang tersebar di beberapa wilayah Indonesia (Kusnandar, 2023). Menurut pengamat pangan Asosiasi Ekonomi Politik Indonesia (AEPI), Pemerintah Indonesia berencana mengimpor satu juta ton beras dari India sebagai bentuk antisipatif guna berjaga-jaga apabila terjadi penurunan produktivitas beras di Indonesia akibat cuaca ekstrim El Nino (Nursyamsi, 2023). Penurunan produksi beras juga dipengaruhi dengan meningkatnya laju pertumbuhan penduduk yang menyebabkan ketersediaan lahan untuk sektor pertanian semakin menyempit, ketersediaan pangan terutama beras semakin sedikit sehingga harga pangan semakin meningkat (Teddy Saputra, 2022). Malthus pernah meramalkan kejadian seperti ini, bahwa jika pertumbuhan pangan kalah saing dengan pertumbuhan penduduk dikhawatirkan akan mengalami kekurangan bahan pangan (Teddy Saputra, 2022). Hal ini menyebabkan ketersediaan beras di Indonesia semakin menurun dan mengakibatkan harga beras yang terus meningkat dalam hitungan hari.

Ketersediaan beras pada tahun 2022 dengan prediksi luas panen padi sekitar 10,45 juta hektar, dibanding tahun 2021 meningkat sebesar 40,87 ribu hektar atau sekitar 0,39% dengan produksi padi sekitar 54,75 juta ton GKG (Gabah Kering Giling). Jumlah produksi beras tahun 2022 sekitar 31,54 juta ton dengan peningkatan produksi sekitar 184,50 ribu ton atau 0,59% dibanding dengan produksi beras tahun 2021 (BPS, 2022). Berdasarkan data yang dihimpun pada laman resmi Badan Pusat Statistik (BPS), data tersebut menjelaskan terkait data produksi padi per tahun pada setiap provinsi di Indonesia. Data tersebut menjelaskan bahwa produksi padi di Indonesia mengalami penurunan dan kenaikan yang berbeda pada setiap tahunnya sesuai dengan wilayah produksi masing-masing.

Dalam penelitian Tri Ifgayani disebutkan bahwa terdapat faktor pendukung yang dapat mempengaruhi produksi padi diantaranya luas lahan, pemakaian benih padi, konsumsi jumlah pupuk, dan jumlah tenaga kerja yang berpengaruh cukup besar terhadap peningkatan produksi padi (Ifgayani et al., 2019). Beberapa faktor tersebut yang berpengaruh secara nyata yaitu faktor jumlah tenaga kerja. Menurut penelitian Nunu Rangga Walis tenaga kerja memiliki korelasi positif namun tidak berpengaruh secara signifikan, dengan koefisien 4,817 berarti setiap penambahan tenaga kerja sebesar 1% produksi padi akan mengalami peningkatan sebesar 4,817% dengan nilai variabel lain tetap (Walis et al., 2021). Friska Erika Moonik menyatakan bahwa dari analisis faktor yang mempengaruhi produksi padi, jumlah tenaga kerja berkorelasi positif dengan koefisien sebesar 0.0310 yang berarti setiap penambahan 1% tenaga kerja, maka produksi padi meningkat sebesar 0.0310 dengan variabel lain bersifat tetap (Moonik et al., 2020). Dalam penelitian Abas mengenai faktor pendukung produksi padi dijelaskan bahwa tenaga kerja dan faktor lainnya mempengaruhi hasil tanaman padi gogo di Kabupaten Morowali, tetapi tidak berpengaruh secara signifikan (Abas et al., 2021). Sedangkan Apriana Rambu Dungu menjelaskan bahwa produksi padi dipengaruhi oleh faktor tenaga kerja sebesar 61,3% dengan koefisien 0,613, dan dipengaruhi faktor lain sebesar 38,7% (Dungu & Retang, 2023). Berdasarkan data tenaga kerja yang dihimpun oleh Badan Pusat Statistik (BPS), data tersebut menjelaskan bahwa jumlah persentase tenaga kerja di sektor pertanian mengalami kenaikan dan penurunan yang berbeda setiap tahunnya. Pengolahan data ini membutuhkan pengolahan dengan teknik

pengelompokan menggunakan *data mining*, karena kapasitas data yang cukup besar. Luaran dari penelitian ini adalah mengelompokkan wilayah dengan tingkat produksi padi tinggi, sedang dan rendah.

Pengelompokan data tersebut menggunakan teknik *data mining*. *Data mining* atau *Knowledge Discovery in Database* (KDD) menurut (Bulolo, 2020) adalah metode yang berhubungan dengan data untuk menemukan informasi dan pola dalam suatu data yang memiliki kapasitas besar. Hasil penggunaan *data mining* dapat digunakan dalam menentukan ketetapan dan memperbaiki pengambilan ketetapan selanjutnya. Dalam *data mining* terdapat beberapa metode seperti dekripsi, klasifikasi (*classification*), prediksi (*prediction*), estimasi (*estimation*), pengklasteran (*clustering*), dan asosiasi (*association*) (Bulolo, 2020).

Metode dalam *data mining* ada beberapa jenis salah satunya adalah *clustering* atau pengelompokan. *Clustering* merupakan metode pengolahan data yang memiliki tujuan untuk membagi objek data kedalam kelompok berdasarkan ciri-ciri yang dimiliki, data yang memiliki ciri sama dikelompokkan dalam kelompok yang sama dari kumpulan data (Supriyatna et al., 2020). *K-Means* adalah salah satu algoritma dalam *clustering*. Algoritma *K-Means* atau algoritma *data mining* bertujuan untuk membagi dan mengelompokkan data dalam beberapa *cluster*. Apabila karakteristik data sama maka akan dikelompokkan dalam kelompok yang sama, jika memiliki perbedaan maka akan dikelompokkan pada kelompok lain (Aulia, 2021). Muhammad Imron menyatakan bahwa menggunakan algoritma *K-Means* dalam mengelompokkan produk di Toko Rizki Barokah dapat mengetahui penetapan jumlah stok produk guna mengurangi penumpukan stok produk (Imron et al., 2020). *K-Means* dalam penelitian Tuslaela dapat digunakan untuk pengelompokan jenis pengolahan dan pendistribusian secara tepat, efektif, dan efisien untuk Program Ketahanan Pangan Jawa Timur 2022 (Tuslaela et al., 2024). Improvisasi algoritma *K-Means* ke dalam sistem intelijensia bisnis dalam penelitian Miwan Kurniawan Hidayat untuk memantau produksi tanaman obat berhasil diterapkan dan menghasilkan kelompok daerah yang potensial dalam menghasilkan tanaman obat, serta memberikan informasi yang ringkas, mudah, cepat dan akurat sesuai kebutuhan pengguna (Hidayat, 2022). Pada tahun 2020 algoritma *K-Means* berhasil diterapkan dalam pengelompokan potensi penyebaran Covid-19 di

Kabupaten Dompu dan *usability* penggunaan aplikasi bekerja dengan sangat baik (Adil et al., 2023). Dalam penelitian Hendro Priyatman penggunaan metode *K-Means* berhasil diterapkan dan membantu dalam mengetahui informasi mengenai prediksi kelulusan mahasiswa (Priyatman et al., 2019).

Berdasarkan pemaparan latar belakang masalah, *K-Means Clustering* digunakan sebagai algoritma pengelompokan pada penelitian ini. Tujuan dalam penelitian ini untuk mengetahui produksi padi pada tiap provinsi di Indonesia dengan kategori tinggi, sedang, dan rendah pada setiap tahunnya. Kemudian dari hasil pengolahan data, diharapkan mampu dijadikan sebagai acuan pemerintah untuk menentukan strategi dalam meningkatkan produksi padi pada daerah yang termasuk dalam kategori rendah.

1.2 Rumusan Masalah

Dalam penelitian merumuskan permasalahan berdasarkan dengan latar belakang yang dipaparkan tersebut yaitu bagaimana penggunaan metode *K-Means Clustering* untuk pengelompokan produksi padi pada provinsi di Indonesia?

1.3 Tujuan

Penelitian ini dilakukan dengan tujuan utama yaitu untuk mengetahui penggunaan metode *K-Means Clustering* untuk pengelompokan produksi padi pada provinsi di Indonesia.

1.4 Batasan Masalah

Beberapa batasan masalah dalam penelitian ini yaitu:

1. Data yang digunakan adalah data sekunder dari Badan Pusat Statistik (BPS) pada tahun 2018 sampai 2022.
2. Data yang diambil adalah data luas panen, produksi, produktivitas padi, dan data persentase tenaga kerja informal sektor pertanian.
3. Pengolahan data yang diperoleh menggunakan algoritma *clustering* yaitu dengan algoritma *K-Means*.

1.5 Manfaat

Terdapat dua manfaat yang diinginkan sesuai dengan penelitian yang dibuat, antara lain sebagai berikut:

1. Manfaat Praktis

Untuk membantu dalam mengetahui daerah potensi produksi padi di Indonesia, sehingga diketahui daerah mana yang memiliki tingkat produksi padi yang baik untuk mencukupi kebutuhan pangan masyarakat Indonesia.

2. Manfaat Teoritis

Perolehan hasil dari penelitian ini dapat dijadikan sebagai sumber rujukan atau referensi untuk penelitian yang akan datang.

