

BAB I

PENDAHULUAN

1.1 Latar Belakang

Buta huruf merupakan suatu keadaan dimana seseorang tidak dapat membaca ataupun menulis, dan biasa diistilahkan dengan istilah tuna aksara (KBBI). Akibat dari tidak meratanya pendidikan akan memberikan dampak besar baik secara langsung atau tidak langsung terhadap wilayah perkotaan dan pedesaan. Luasnya jangkauan wilayah dan infrastruktur yang belum cukup memadai, serta sumber daya manusia yang kurang menyebabkan ketidakmerataan pendidikan belum dapat menyeluruh dengan baik. Hal ini dapat dilihat pada wilayah - wilayah atau daerah - daerah pinggiran dan terpencil. Di lain pihak, tidak meratanya pendidikan inilah yang membuat kualitas pendidikan terfokus pada satu wilayah saja dan akhirnya hal tersebut dapat berdampak pada tingginya jumlah buta huruf pada wilayah-wilayah lainnya.

Permasalahan buta huruf ini menjadi sebuah tantangan yang cukup besar bagi dunia pendidikan Indonesia untuk dapat diselesaikan bersama. Menurut data Badan Pusat Statistik, angka buta huruf usia 15 keatas (15+, 15-44, dan 45+) tertinggi dari tahun 2015 sampai tahun 2019 berada pada provinsi Papua, yaitu pada tahun 2015 total jumlah persentase dari penduduk buta huruf setinggi 89,21%, pada tahun 2016 setinggi 88,6%, pada tahun 2017 setinggi 81,23%, pada tahun 2018 setinggi 73,43%, pada tahun 2019 setinggi 69,32%. Hal ini berdasarkan data persentase penduduk buta huruf, namun jika dilihat dari data proyeksi penduduk menurut provinsi (BPS), provinsi Papua pada tahun 2019 berada di posisi ke-6 dengan jumlah sekitar 2183,16 (ribu jiwa) dan provinsi Jawa Timur berada di posisi tertinggi dengan jumlah sekitar 9840,09 (ribu jiwa). Banyaknya penduduk yang masih buta huruf merupakan salah satu dampak tingginya tingkat kemiskinan yang akhirnya menjadikan penduduk atau masyarakat tidak mampu atau tidak memiliki kesempatan untuk dapat mengenyam bangku pendidikan. Agar dapat

menurunkan tingkat buta huruf di Indonesia, diperlukan upaya bersama dalam merangkul atau mengajak masyarakat sesuai dengan karakteristik dan budaya disetiap daerah-daerah, sehingga dapat menumbuhkan motivasi serta minat baca masyarakat.

Sebelumnya, penelitian telah dilakukan oleh (Warnia & Fadly. A, 2017) dengan judul “Klasterisasi Tingkat Buta Huruf di Indonesia berbasis Point-Based K-Means Analysis” menggunakan dataset pada tahun 2015 berupa data persentase sejumlah 102 data yang terbagi menjadi 3 *cluster* berdasarkan tingkat usia (15+, 15-44, dan 45+) dengan menggunakan *rapid miner* sebagai *tool* untuk mengolah data, serta *point-based* digunakan untuk mengetahui provinsi mana yang memiliki tingkat buta huruf tertinggi. Pada penelitian tersebut, peneliti tidak menggunakan pengukuran *cluster* terbaik. Hasil dari penelitian tersebut didapat provinsi Papua memiliki tingkat buta huruf tertinggi pada *cluster* tingkat usia 45 keatas sebesar 31,57% yang terdapat pada tingkat usia 45+.

Pada penelitian yang lain dengan judul “Analisa Segmentasi Pelanggan Dengan RFM Model Pada PT. Arthamas Citra Mandiri Menggunakan Metode *Fuzzy C-Means Clustering*” menggunakan algoritma *Fuzzy C-Means* yang bertujuan untuk mengelompokkan pelanggan berdasarkan data riwayat transaksi sebanyak 981 transaksi dengan periode selama 12 bulan di tahun 2017 agar dapat mengetahui pelanggan mana yang menguntungkan bagi perusahaan (Taqwim et al., 2019). Hasil yang didapat dari 4 *cluster*, *cluster* 3 merupakan hasil terbaik dengan nilai SSE (*Sum of Squares Error*) yaitu 0,6296328. *Cluster* optimum yang didapat pada penelitian segmentasi *customer* berada pada 4 *cluster* dengan skenario 2 *cluster* sampai 10 *cluster* . Karena buta huruf masih merupakan salah satu permasalahan yang belum terselesaikan, maka penelitian ini menggunakan data buta huruf dari tahun 2015 sampai dengan tahun 2021 di 33 provinsi di Indonesia dengan mengelompokkannya untuk mengetahui ada atau tidaknya kemajuan atau perubahan yang terjadi antara tahun 2015 sampai dengan tahun 2021 pada masing-masing provinsi. Perubahan yang dimaksud dapat berupa kenaikan

dan penurunan atau tidak ada perubahan pada jumlah penduduk yang mengalami buta huruf. Oleh karena itu, studi ini diharapkan dapat menjadi metode alternatif untuk mengkategorikan data angka buta huruf sehingga pembaca dan peneliti lainnya dapat menyadari pentingnya permasalahan tingginya jumlah penduduk buta huruf yang dihadapi.

Pada penelitian lain dengan judul “Perbandingan Algoritma *K-Means* dengan Algoritma *Fuzzy C-means* untuk *Clustering* Tingkat Kedisiplinan Kinerja Karyawan” menggunakan dua algoritma yaitu *Fuzzy C-Means* dan *K-Means* untuk dibandingkan, peneliti bertujuan untuk mengetahui algoritma manakah yang lebih baik dalam melakukan metode *clustering*. Berdasarkan hasil penelitian tersebut dapat diketahui bahwa algoritma *Fuzzy C-Means* lebih baik dalam melakukan *clustering* atau pengelompokan dari pada algoritma *K-Means*, karena hasil validitas algoritma *Fuzzy C-Means* memiliki nilai yang lebih optimal sebesar 0,758 dikarenakan nilai validitas lebih besar atau mendekati nilai 1, sedangkan algoritma *K-Means* hanya memiliki nilai validitas sebesar 0,528 yang menunjukkan bahwa nilai validitas algoritma *K-Means* lebih kecil dan tidak mendekati nilai 1 jika dibandingkan dengan nilai validitas algoritma *Fuzzy C-Means* (Agustina & Prihandoko, 2018), (Ramadhan et al., 2017), (Putri & Dwidayati, 2021).

Oleh karena itu, *Fuzzy C-Means* yang merupakan salah satu pengelompokan data yang derajat keanggotaannya menentukan setiap titik data dalam sebuah cluster, digunakan dalam penelitian ini, selain itu pada penelitian lainnya dengan data yang hampir serupa telah dilakukan pengelompokan dengan metode *K-Means*. Metode *Partition Coefficient Index* (PCI) dipilih untuk menghasilkan informasi yang diinginkan serta mudah untuk diimplementasikan, yaitu dengan cara melihat *cluster* dengan nilai validitas tertinggi. Maka penelitian ini memilih menggunakan *fuzzy c-means* dan metode *Partition Coefficient Index* (PCI) karena pada penelitian sebelumnya dengan dataset yang sama telah terlebih dahulu diolah menggunakan algoritma *k-means*. Maka, penelitian ini dibuat dengan judul

“Implementasi Algoritma *Fuzzy C-Means* untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Tingkat Buta Huruf”.

1.2 Rumusan Masalah

Diantara permasalahan yang muncul dan dapat diambil dari latar belakang di atas, yaitu:

1. Berapa *cluster* optimum yang terdapat pada pengelompokan tingkat buta huruf menggunakan metode *Partition Coefficient Index* pada *Fuzzy C-Means*?
2. Berapa provinsi di Indonesia yang termasuk dalam setiap *cluster* optimum?

1.3 Batasan Penelitian

Adapun beberapa batasan masalah yang terdapat pada penelitian ini yang perlu diketahui, antara lain:

1. Data yang digunakan merupakan data dari BPS yaitu “Persentase penduduk buta huruf menurut kelompok umur” dan akan melalui pra proses terlebih dahulu.
2. Data yang digunakan terdiri dari 33 provinsi periode tahun 2015 sampai dengan tahun 2021 sebanyak 231 data.
3. *Tools* yang digunakan untuk *clustering* adalah perangkat lunak R Studio.
4. Penentuan *cluster* optimum berdasarkan 2 *cluster* sampai 10 *cluster*.

1.4 Tujuan Penelitian

Adapun tujuan penelitian yang ingin dicapai berdasarkan pada rumusan masalah di atas adalah:

1. Untuk menentukan jumlah *cluster* optimum pada tingkat buta huruf di 33 provinsi di Indonesia.
2. Untuk mengetahui kelompok provinsi di Indonesia yang termasuk pada setiap *cluster* yang optimum.

1.5 Manfaat Penelitian

Adapun manfaat yang dapat diperoleh dari penelitian ini yaitu:

1. Menjadi salah satu alternatif penelitian untuk penerapan *clustrering* data pada tingkat buta huruf di Indonesia.
2. Membantu pembaca dan peneliti selanjutnya mengetahui provinsi-provinsi yang memiliki jumlah tingkat buta huruf tertinggi.
3. Menjadi salah satu rujukan untuk meningkatkan kesadaran pembaca dan peneliti selanjutnya akan pentingnya penanganan buta huruf.

