

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Data mining adalah sebuah proses menemukan pola dan tren yang bermakna dengan menggunakan teknologi sosialisasi pola serta statistik dan teknik matematika (Larose, 2006). Salah satu metode dalam data mining adalah *K-Nearest Neighbor* (Sigit, 2018). Metode *K-Nearest Neighbor* merupakan teknik prediksi pengelompokan dalam satu *record* dengan mencari nilai prediktor yang sama di dalam basis data histori dan nilai prediksi dari *record* yang terdekat (Indrajani, 2015). Metode *K-Nearest Neighbor* bekerja dengan konsep bahwa data akan memiliki kelas dengan data sekitarnya, konsep ini disebut konsep ketetangaan. *K-Nearest Neighbor* adalah metode yang menerapkan hasil yang tegas, hanya menghasilkan keputusan Ya atau Tidak.

Fuzzy K-Nearest Neighbor atau (FK-NN) merupakan pengembangan dari *K-Nearest Neighbor* atau (K-NN) yang digabung dengan logika *fuzzy* (Prasetyo, 2012). Logika *fuzzy* merupakan logika yang memiliki elemen ketidakpastian (kabur), dimana akan menghasilkan rentang nilai antara Ya dan Tidak. *Fuzzy K-Nearest Neighbor* akan menghasilkan prediksi yang jauh lebih deskriptif dibandingkan Metode *K-Nearest Neighbor*, karena hasil prediksinya tidak hanya Ya atau Tidak namun hasil prediksi menunjukkan berapa nilai untuk Ya dan Tidak. Diharapkan dengan menggunakan metode pengembangan *K-Nearest Neighbor* yaitu *Fuzzy K-Nearest Neighbor* dalam penelitian ini akan memberikan gambaran prediksi secara deskriptif dan detail, sehingga mengetahui berapa nilai untuk Ya dan berapa nilai untuk Tidak.

Corona Virus Disease 2019 atau disingkat sebagai Covid 19 adalah penyakit yang disebabkan oleh virus SARS-COV2 dan menyerang sistem pernapasan manusia sehingga dikategorikan sebagai penyakit berbahaya

karena dapat merenggut nyawa penderitanya (Sutejo, 2020). Kasus pertama Covid 19 tercatat di Wuhan China pada tanggal 30 Desember 2019 dan diduga penularan virus ini berasal dari kelelawar kemudian berpindah ke manusia. Virus tersebut terkonfirmasi dapat menular antar manusia dan penularannya menyerupai penyakit flu dan batuk. Menurut (*World Health Organization, 2020*), penderita virus ini ketika bernafas ataupun batuk akan mengeluarkan beberapa tetes cairan yang mengandung virus. Hal tersebut yang menjadikan mudahnya penyebaran virus tersebut hingga pada 11 Maret 2020 Covid 19 ditetapkan sebagai penyakit pandemi oleh *World Health Organization*. Menurut KBBI pandemi adalah wabah yang berjangkit serempak dimana-mana atau meliputi geografi yang luas. Penyebaran Covid 19 di seluruh dunia mencatat kasus positif pada tanggal 16 Januari 2023 sebesar 663,001,898 (*World Health Organization, 2023*). Covid 19 juga menyebar sampai di Indonesia. Tercatat kasus pertama Covid 19 di Indonesia pada Maret 2020.

Hampir seluruh negara di dunia mengalami dampak akibat adanya virus ini, termasuk Indonesia. Tercatat kasus positif penyebaran Covid 19 di Indonesia pada tanggal 19 Januari 2023 sebesar 6.727.317 (*Covid19.go.id, 2023*). Gejala-gejala Covid 19 yaitu demam, batuk, sakit tenggorokan, malaise, sakit kepala, nyeri otot, mual, muntah, diare, dan kehilangan penciuman (Makarim, 2021). Berdasarkan tingkat keparahan kasus gejala Covid 19 dibagi menjadi 5 yaitu : tanpa gejala, gejala ringan, gejala sedang, gejala berat, dan gejala kritis. Tingkat keparahan dapat dipengaruhi oleh beberapa faktor, yaitu usia, jenis kelamin, dan penyakit komorbid (Febriana, 2020). Komorbid adalah kondisi dimana seseorang memiliki dua atau lebih masalah kesehatan pada saat yang sama (Christiansen, 2021). Per tanggal 22 Januari 2023 kasus kematian akibat Covid 19 di dunia sebesar 6,713,093 dan di Indonesia sebesar 160,781. Menurut situs (*Covid.go.id, 2022*), tingkat kepulihan dari Covid 19 sangatlah tinggi sekitar 80% dapat pulih sendiri, tanpa adanya perawatan khusus. Ditambah vaksin Covid 19 yang saat ini dilakukan sebanyak 4 dosis (*Covid.go.id, 2023*). Terbukti bahwa vaksin Covid 19 mampu melindungi diri dari paparan virus (Younus, 2021), sebagai upaya tambahan dalam menekan angka kasus kematian akibat Covid 19. Namun metode perawatan setiap pasien juga harus diperhatikan. Pada Juli 2021 terjadi kondisi dimana pasien yang membutuhkan penanganan prioritas tidak bisa diterima oleh rumah sakit dikarenakan penuhnya rumah sakit (CNN, 2021).

Kasus penyebaran konfirmasi positif dan angka kematian akibat Covid 19 saat ini sudah jauh menurun, namun tidak menutup kemungkinan terjadi kembali lonjakan kasus konfirmasi positif Covid 19 dan juga angka kematian akibat virus tersebut. Hal tersebut diperkirakan akibat varian Covid 19 yang terus bermutasi sehingga besar kemungkinan akan adanya varian virus Covid 19 yang berbeda. Sebagai upaya pencegahan agar tidak terjadi pasien yang tidak mendapatkan penanganan prioritas, membutuhkan analisis berbasis data yang dapat memprediksi tingkat harapan kesembuhan pasien berdasarkan faktor usia, jenis kelamin, penyakit komorbid, dan status vaksinasi pasien. Diharapkan hasil prediksi yang baik dapat memberikan gambaran status tingkat harapan kesembuhan pasien, sehingga pasien yang memiliki keparahan tinggi menjadi prioritas penanganan oleh petugas medis. Dalam menentukan keputusan pasien mana yang membutuhkan perawatan prioritas, data mining dibutuhkan dalam analisis ini.

Pada penelitian Pasien Covid 19 sebelumnya yang menggunakan data mining dilakukan oleh (Lilia, dkk, 2021) yang menggunakan Metode *Naive Bayes* untuk mengklasifikasikan tingkat risiko pasien Covid 19, (Rahmi, dkk, 2023) yang menggunakan Metode *Neural Network System* dan *Logistic Regression* untuk mengklasifikasikan gejala Covid 19, (Mustaghfiroh, dkk, 2022) yang menggunakan Metode *K-Nearest Neighbor* untuk mengklasifikasikan Pasien Covid 19 di Indonesia, (Raka, dkk, 2022) yang menggunakan Metode *K-Nearest Neighbor* untuk mengklasifikasikan Pasien Covid 19 yang membutuhkan perawatan di ICU, dan (Candra, dkk, 2022) Metode *Naive Bayes* untuk memprediksi tingkat kesembuhan Pasien Covid 19. Dalam skripsi ini akan digunakan Metode *K-Nearest Neighbor*.

Alasan memilih metode ini karena *K-Nearest Neighbor* memiliki keakuratan yang tinggi dibandingkan metode klasifikasi yang lain. Penelitian yang dilakukan (Sulastri, dkk, 2019) membandingkan 3 metode klasifikasi yaitu *K-Nearest Neighbor*, *Naive Bayes* dan *Neural Network System*. Hasil keakuratan dari 3 metode tersebut *K-Nearest Neighbor* memiliki keakuratan paling tinggi dibanding 2 metode tersebut. Pada penelitian sebelumnya yang menggunakan Metode *K-Nearest Neighbor* untuk memprediksi pasien Covid 19, hasil prediksi dalam penelitian tersebut menghasilkan hasil tegas yang hanya menunjukkan apakah pasien tersebut sembuh atau meninggal. Sebagai pengembangan penelitian dibutuhkan hasil prediksi yang jauh lebih deskriptif dalam memprediksi pasien Covid 19 sehingga hasil yang di

tunjukan dapat memberikan gambaran pasien mana yang jauh lebih diprioritaskan penanganannya. Oleh karena itu penggabungan logika *fuzzy* dengan Metode *K-Nearest Neighbor* diharapkan akan menghasilkan prediksi Pasien Covid 19 yang lebih deskriptif.

Penelitian ini bertujuan untuk memprediksi tingkat harapan kesembuhan pasien dengan otomatis dan cepat berdasarkan usia, penyakit komorbid, jenis kelamin, dan status vaksinasi. Pemilihan Metode *Fuzzy K-Nearest Neighbor* diharapkan dapat menghasilkan klasifikasi dan prediksi yang memiliki ketepatan akurasi yang tinggi dan deskriptif, sehingga penanganan terhadap Pasien Covid 19 jauh lebih tepat sasaran. Berdasarkan penjabaran latar belakang tersebut, maka penelitian ini diberi judul :
“IMPLEMENTASI METODE FUZZY K-NEAREST NEIGHBOR DALAM MEMPREDIKSI TINGKAT HARAPAN KESEMBUHAN PASIEN AKIBAT COVID 19 “

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan dapat dirumuskan permasalahan sebagai berikut :

1. Bagaimana implementasi Metode *Fuzzy K-Nearest Neighbor* dalam memprediksi tingkat harapan kesembuhan pasien?
2. Bagaimana efektivitas Metode *Fuzzy K-Nearest Neighbor* dalam tingkat harapan kesembuhan pasien berdasarkan jumlah K atau jumlah tetangga terdekat yang digunakan dalam memprediksi?
3. Bagaimana efektivitas penggunaan nilai p atau parameter yang mengontrol bentuk jarak yang berbeda-beda pada Jarak Minkowski untuk Metode *Fuzzy K-Nearest Neighbor* dalam memprediksi tingkat harapan kesembuhan pasien?
4. Bagaimana implementasi Metode *Fuzzy K-Nearest Neighbor* pada bahasa pemrograman R?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah di sampaikan maka tujuan dari penulis sebagai berikut :

1. Mengetahui bagaimana Metode *Fuzzy K-Nearest Neighbor* dalam memprediksi tingkat kesembuhan pasien.
2. Mengetahui bagaimana efektivitas *Fuzzy K-Nearest Neighbor* dalam memprediksi tingkat harapan kesembuhan pasien berdasarkan jumlah K yang digunakan.
3. Mengetahui bagaimana efektivitas Jarak Minkowski berdasarkan penggunaan nilai p pada Metode *Fuzzy K-Nearest Neighbor* dalam memprediksi tingkat harapan kesembuhan pasien.
4. Mengetahui bagaimana implementasi Metode *Fuzzy K-Nearest Neighbor* pada bahasa pemrograman R.

1.4 Batasan Masalah

Penelitian ini akan membahas batasan masalah utama yang telah ditetapkan antara lainnya sebagai berikut :

1. Pengambilan data sekunder terkait pasien Covid 19 yang dirawat pada tahun 2022 yang diperoleh dari Puskesmas Cempaka Putih.
2. Sampel yang diambil menggunakan Teknik *Disproportionate Stratified Random Sampling* didapatkan sampel sebanyak 375 pasien Covid 19 yang dirawat pada tahun 2022.
3. Variabel Prediktor yang digunakan adalah Usia, Jenis Kelamin, Diabetes Melitus, Hypertensi, Kardiovaskular, dan Status Vaksinasi (dosis 1,2,dan 3)
4. Penelitian ini menggunakan Metode *K-Nearest Neighbor* dengan menggunakan logika *fuzzy*.

1.5 Manfaat Penelitian

Penelitian ini memberikan manfaat praktis dan pemahaman mendalam terhadap permasalahan yang dibahas antara lainnya sebagai berikut:

1. Bagi ilmu pengetahuan, sebagai wawasan tambahan terkait Metode *Fuzzy K-Nearest Neighbor* dalam memprediksi tingkat harapan kesembuhan dan tingkat risiko kematian akibat Covid 19.
2. Bagi pelayanan kesehatan, sebagai referensi dalam menentukan perawatan pasien yang memiliki prioritas tinggi.