

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kesehatan merupakan salah satu faktor terpenting bagi makhluk hidup khususnya manusia, karena tanpa kesehatan yang baik, makhluk hidup maupun manusia akan kesulitan untuk melakukan aktivitas sehari-hari. Kesehatan adalah keadaan fisik, mental, spiritual dan sosial tubuh yang memungkinkan individu untuk hidup produktif secara sosial dan ekonomi. Dari segi kesehatan, terdapat pula berbagai kondisi terkait masalah lingkungan alam dan buatan yang mempengaruhi hasil (dampak) kesehatan. Penyakit ini, baik yang dapat dicegah maupun yang dapat disembuhkan, tidak dapat dihindari, salah satunya adalah penyakit ginjal kronik (PGK)[1].

Masalah yang sering dihadapi oleh masyarakat yaitu besarnya biaya perawatan terkait penyakit ginjal kronik, perawatan tersebut kurang bisa memberikan hasil dan belum adanya penawar untuk menyembuhkan penyakit tersebut. Di negara ini sendiri yang telah terdata berkisar 70.000 pasien pengidap ginjal kronik. survei dari WHO di negara ini mengalami lonjakan penderita ginjal kronik sebanyak 46% dari tahun 1955 hingga tiga tahun kedepan[2].

Penyakit ginjal kronik (PGK) adalah penyakit dimana ginjal penderita tidak bekerja dengan baik, bila ginjal tidak bekerja secara efektif maka akan mengakibatkan kerusakan ginjal. Seringkali pasien tidak menyadari gejala pada awal penyakit. Dengan demikian, pasien hanya mengetahui bahwa penyakit ginjal kronik telah didiagnosis melalui tes darah sehingga banyak indikasi mengarah ke Penyakit ginjal kronik. Diagnosis dibuat karena faktor risiko untuk kondisi lain seperti diabetes serta tekanan darah tinggi. Untuk menentukan PGK, perlu untuk mengklasifikasikan dengan bantuan indikasi yang berbeda yang menjadi standar acuan dalam diagnosis PGK[3].

Pada klasifikasi penyakit tersebut telah dilakukan penelitian sebelumnya dengan berbagai metode klasifikasi diantaranya, Metode ini digunakan pada penelitian terdahulu yang diteliti oleh Warid Yunus menggunakan data *chronic kidney disease* yang diperoleh dari UCI dengan metode KNN serta KNN ditambah

PSO hasil akurasi KNN sebesar 78,75%, selain itu hasil metode dengan penambahan PSO memiliki akurasi 97,25%[1]. Penelitian lebih lanjut oleh Hilda Amalia dengan data yang diperoleh dari situs UCI dan membandingkan metode SVM dan NN, diperoleh nilai akurasi NN 93,36% dan nilai akurasi SVM 95,16%[2]. Penelitian tambahan dilakukan oleh Ardina Ariani dan Samsuyardi dengan data yang diperoleh dari situs UCI menggunakan metode *K-Nearest Neighbor* (KNN) menghasilkan akurasi yang baik yaitu akurasi sebesar 85,83%[3]. Penelitian lebih lanjut oleh Samsilul Azhar, dan kawan – kawan data dari rumah sakit Dr. M.Yunus bengkulu dengan observasi menggunakan sistem pakar proses pengolahan data dan proses konsultasi dilakukan dengan cepat serta menghasilkan laporan yang cukup akurat [4]. Penelitian lebih lanjut oleh Andi Sri Irtawaty dengan proses pengumpulan data primer (data pakar) dari beberapa literatur metode *k-means* mempunyai akurasi 90%[5]. Penelitian selanjutnya oleh Sutri Handayani data yang diperoleh adalah data sekunder karena diperoleh dari *UCI repository* akurasi 95.00 % dengan menggunakan algoritma SVM, 97.25% menggunakan algoritma C4.5[6].

Dalam solusi yang ada memiliki kelebihan tersendiri akan tetapi semua metode memiliki kekurangan yang hampir sama yaitu bermasalah pada besar data. Data tersendiri beragam bentuk dan model sehingga akan mempengaruhi keefektifan metode. Sehingga hasil akurasi setiap metode akan berbeda meskipun menggunakan data sama[7],.

Untuk mempermudah pendeteksian Penyakit ginjal kronik (PGK) dapat digunakan berbagai macam metode klasifikasi, penelitian ini memakai SVM (*Support Vector Machine*) dalam proses klasifikasi[8].

SVM merupakan salah satu metode prediksi, baik dalam kasus klasifikasi yang bertujuan untuk mencari hyperplane yang terbaik untuk memisahkan dua class yang terdapat didalam input space[9].

Beberapa keuntungan yang ada pada SVM diantaranya[7]:

1. *Curse of dimensionality*, khususnya, adalah permasalahan yang sering ditemui oleh metode pengenalan pola yang terdapat pada estimasi parameter. Karena sampel data yang kemungkinan kecil di komparasi dengan ukuran ruang data vektor, maka besar pula ukuran ruang vektor

yang akan diproses, maka semakin banyak pula jumlah data yang dibutuhkan dalam dimensi yang lebih tinggi.

2. *Generalisasi*, yaitu ketika metode (SVM, *neural network*, dll) memiliki kemampuan untuk mengklasifikasikan ke dalam model dan data yang digunakan dalam fase pembelajaran tidak termasuk dalam metode.
3. *Feasibility*. Metode SVM dapat diimplementasikan dikarenakan, penentuan support vector bisa dibangun ke QP problem.

Atas dasar penjelasan tersebut, penulis meneliti dengan judul “Klasifikasi Penyakit ginjal kronik menggunakan *Support Vector Machine*”. Dalam penelitian ini, peneliti mengukur akurasi menggunakan metode SVM. Penelitian ini menghasilkan metode yang paling efektif dalam klasifikasi penyakit ginjal.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Pengidap PGK sulit diidentifikasi sampai menyebabkan meninggal dunia dikarenakan klasifikasi yang kurang akurat. Proses klasifikasi penyakit ginjal yang membutuhkan suatu metode dengan performa yang baik agar hasil akurasi yang dihasilkan tinggi.

1.2.2 Metode Usulan

Dari permasalahan tersebut akan dilakukan penelitian untuk mengklasifikasi Penyakit ginjal kronik menggunakan metode SVM,

1.2.3 Pertanyaan Penelitian

Berapa nilai akurasi metode SVM dalam pembagian data menggunakan *k-fold* dan manual untuk proses klasifikasi Penyakit ginjal kronik?

1.3 Tujuan dan Manfaat Penelitian

1.3.1 Tujuan Penelitian

Mengetahui akurasi yang akan dihasilkan oleh metode SVM dalam klasifikasi Penyakit ginjal kronik.

1.3.2 Manfaat Penelitian

Manfaat pada penelitian kali ini diharapkan dapat mengetahui nilai akurasi dari metode SVM. Juga untuk membantu para ahli mendiagnosis Penyakit ginjal kronik.

1.4 Batasan – batasan

Mengenai batasan masalah dari penelitian ini yakni:

1. Sistem yang dibuat menggunakan Bahasa pemrograman *Python* yang ada pada *google colaboratory*.
2. Data yang digunakan sebanyak 400 data dengan jumlah fitur 24 yaitu, *age, bp, sg, al, su, rbc, pc, pcc, ba, bgr, bu, sc, sod, pot, hemo, pcv, wc, rc, htn, dm, cad, appet, pe, ane*, dan 2 label yaitu 1 dan 0, 1 berarti positif Penyakit ginjal kronik (PGK), sedangkan 0 berarti negatif Penyakit ginjal kronik (PGK). data ini diperoleh dari Kaggle.com(<https://www.kaggle.com/akshayksingh/kidney-disease-dataset>).

1.5 Sistematika Laporan

Mengenai sistematika yang digunakan dalam penulisan skripsi terdiri dari beberapa bab, yaitu:

BAB I PENDAHULUAN

Terkait dalam bab ini akan menjelaskan latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, serta sistematika penulisan laporan.

BAB II KAJIAN *LIBRARY*

Pada bab akan menjelaskan tentang konsep dan teori yang mendukung dalam penyusunan penelitian ini. Serta landasan teori dan penelitian sebelumnya yang akan digunakan dalam mengatasi permasalahan.

BAB III METODE USULAN

Pada bab akan menjelaskan terkait dengan rancangan sistem yang akan dibangun, tertera adanya data penelitian, arsitektur sistem, *flowchart* metode SVM, tahapan penelitian, serta pengujian yang digunakan.

BAB IV HASIL dan PEMBAHASAN

Pada bagian ini menjelaskan tentang implementasi sistem yang dibuat berdasarkan hasil rancangan sebelumnya beserta hasil pengujian.

BAB V PENUTUP

Pada bab ini berisi kesimpulan dari implementasi sistem yang telah dibuat serta saran untuk pengembangan sistem.