

BAB I PENDAHULUAN

Bab I memuat latar belakang dari topik penelitian, rumusan masalah, tujuan penelitian, batasan penelitian, serta manfaat dari penelitian. Kemudian pada akhir bab ini dibahas mengenai sistematika penulisan yang menjadi acuan penulis dalam penyusunan isi dari karya ilmiah ini.

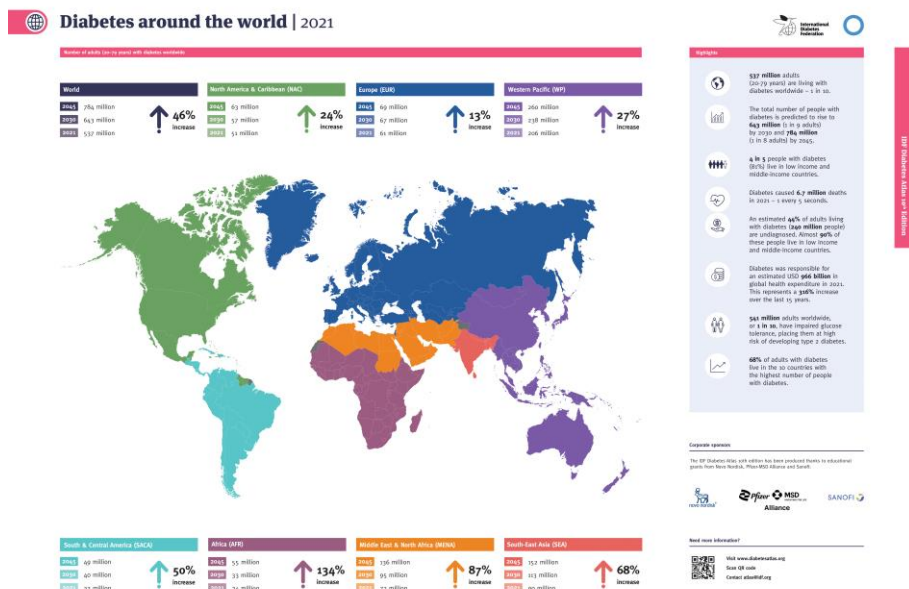
I.1 Latar Belakang

Selama beberapa dekade terakhir, isu kesehatan menjadi salah satu fokus utama di seluruh dunia. Banyak tersebarnya penyakit menular dan tidak menular, penyakit baru maupun penyakit lama, serta akses pelayanan kesehatan yang berkualitas menjadi perhatian utama dalam upaya meningkatkan kualitas hidup masyarakat. Di Indonesia, regulasi terkait kesehatan telah diatur di dalam undang-undang, salah satunya yaitu Undang - Undang Kesehatan Republik Indonesia No. 36 Tahun 2009. Di dalam Undang-undang tersebut dijelaskan bahwa kesehatan adalah keadaan sejahtera yaitu baik kesehatan badan, jiwa, maupun sosial yang memungkinkan setiap orang hidup produktif baik secara sosial dan ekonomis (Menteri Hukum dan Hak Asasi Manusia Republik Indonesia, 2009). Kesehatan yang optimal adalah hak setiap individu, dan pemeliharaan catatan medis yang akurat serta terpercaya adalah salah satu kunci dalam upaya mencapai standar pelayanan kesehatan yang bermutu. Dalam konteks ini, penyakit menjadi tantangan utama yang perlu diatasi, sehingga pemahaman mendalam tentang penyakit juga menjadi unsur penting dalam memberikan pelayanan kesehatan yang berkualitas. Salah satu tantangan utama dalam dunia kesehatan yang mendapat sorotan khusus adalah penyakit diabetes.

Diabetes yaitu penyakit berupa metabolik kronis, dengan disertai tanda-tanda seperti peningkatan kadar glukosa darah. Seiring berjalannya waktu, penyakit ini akan menyebabkan kerusakan serius pada beberapa organ tubuh seperti jantung, pembuluh darah, mata, ginjal, dan saraf (World Health Organization, 2023). Semua jenis diabetes dapat menyebabkan komplikasi di banyak bagian tubuh dan meningkatkan risiko kematian dini secara keseluruhan seperti serangan jantung,

stroke, gagal ginjal, amputasi, kehilangan penglihatan, dan kerusakan saraf (World Health Organization, 2021).

Tahun 2014, 8.5% dari populasi dewasa yang berusia 18 tahun ke atas mengalami diabetes. Kemudian pada tahun 2019, diabetes menjadi penyebab langsung dari 1.5 juta kematian, dengan 48% dari total kematian akibat diabetes terjadi sebelum mencapai usia 70 tahun. Sekitar 460000 kematian yang disebabkan oleh penyakit ginjal lainnya terkait dengan diabetes, dan peningkatan kadar glukosa darah juga berkontribusi sebanyak 20% terhadap kematian kardiovaskular. Antara tahun 2000 hingga 2019, terjadi peningkatan sekitar 3% dalam angka kematian akibat diabetes jika diukur dengan standar usia. Pada negara-negara berpendapatan menengah ke bawah, tercatat peningkatan sebesar 13% dalam angka kematian akibat diabetes (World Health Organization, 2023). Berikut merupakan gambar persebaran penyakit diabetes yang diterbitkan oleh *International Diabetes Federation (IDF)*.



Gambar I.1 Data Diabetes Around the World

Sumber: (Alberti, 2021)

Gambar I.1 menunjukkan pada akhir tahun 2021, *International Diabetes Federation (IDF)* dalam edisi ke-10 Atlas mengonfirmasi bahwa diabetes merupakan salah satu kegawatdaruratan kesehatan global yang mengalami

pertumbuhan paling cepat di abad ke-21 ini. Pada tahun 2021, lebih dari setengah miliar orang di seluruh dunia hidup dengan diabetes, atau tepatnya mencapai 537 juta individu. Proyeksi untuk tahun 2030 menunjukkan angka ini diperkirakan akan mencapai 643 juta, dan meningkat menjadi 783 juta pada tahun 2045 (Alberti, 2021). Dengan prevalensi global diabetes yang terus meningkat, situasi di Indonesia juga menunjukkan dampak serius dari masalah kesehatan ini, memperkuat urgensi untuk memahami dan mengatasi tantangan diabetes di tingkat nasional. Mengingat dampak serius yang diakibatkan oleh diabetes, peringatan dini (*early warning*) terhadap penyakit ini sangat diperlukan. Oleh karena itu, penyakit diabetes memerlukan pendekatan holistik dalam pencegahan, diagnosis, dan pengolahannya. Peringatan dini memungkinkan tindakan pencegahan dan penanganan lebih cepat, yang dapat mencegah komplikasi serius. Salah satu caranya adalah dengan mengetahui gejala awal atau mendapatkan diagnosis yang akurat. Disinilah teknologi *data mining* dapat memainkan peran penting yang berfungsi untuk memperkuat hasil diagnosis berupa prediksi, meskipun prediksi ini bukan sebagai rujukan diagnosis utama.

Penelitian terdahulu yang berjudul “*Augmentasi Awal untuk Diagnosis Diabetes Menggunakan Pendekatan Data Mining*” oleh Ginting (2021) menunjukkan bahwa teknik *data mining* dapat membantu meningkatkan diagnosis diabetes. *Data mining* sering diartikan sebagai *knowledge discovery in database* atau disingkat menjadi KDD yang erat kaitannya dengan *machine learning* (Pahlevi, 2018). *Machine learning* merupakan inti dari *data mining* dalam hal pengembangan model prediksi dan pengenalan pola. Terdapat beberapa algoritma *machine learning* yang termasuk ke dalam jenis-jenis model *data mining* seperti *Support Vector Machine*, *Random Forest*, *Neural Network*, *K-Nearest Neighbors*, *Naïve Bayes*, *Gradient Boosting Machine*, dan lain-lain. Algoritma *Gradient Boosting Machine* sebagai salah satu model *data mining* adalah teknik yang dapat meningkatkan kinerja model secara rekursif, menyesuaikan pembelajar yang lemah dengan yang tersisa, menemukan struktur dan hubungan data yang kompleks (Zhang dkk., 2019).

Penelitian dari Zhang dkk., (2019) yang berjudul “*Predictive Analytics with Gradient Boosting in Clinical Medicine*” menunjukkan bahwa algoritma *Gradient*

Boosting Machine memiliki kemampuan yang baik dalam menangani data yang kompleks dan beragam, terutama dalam bidang kedokteran klinis. Algoritma *Gradient Boosting Machine* dapat mengidentifikasi hubungan non-linear antara berbagai variabel klinis dan hasil kesehatan pasien, yang seringkali sulit dideteksi dengan metode tradisional. Penelitian ini juga menemukan bahwa algoritma *Gradient Boosting Machine* memberikan evaluasi yang lebih tinggi dibandingkan dengan metode lainnya dilihat dari nilai AUC berdasarkan kurva ROC, yang menjadikannya alat yang efektif untuk analisis prediktif dalam konteks medis. Hal ini memperkuat keputusan untuk menggunakan algoritma *Gradient Boosting Machine* dalam penelitian ini, karena keunggulannya dalam memproses dan menganalisis data yang kompleks dapat meningkatkan keakuratan prediksi diabetes melitus berdasarkan data rekam medis yang tersedia.

Algoritma *Gradient Boosting Machine* juga sudah diimplementasikan pada proses pemeriksaan status kesehatan janin yang dilakukan oleh Olayemi & Olasehinde, (2024) dalam penelitian yang berjudul “*Machine Learning Prediction of Fetal Health Status from Cardiotocography Examination in Developing Healthcare Contexts*”. Penelitian tersebut menunjukkan kinerja yang sangat baik dari algoritma *Gradient Boosting Machine* dengan berbagai metrik evaluasi seperti akurasi sebesar 0.8009 dan hasil *f-1 score* 0.8665, yang menunjukkan keseimbangan antara *presisi* dan *recall* dalam prediksi. Hasil tersebut menginterpretasikan prediksi kondisi janin dengan akurat, yang sangat penting untuk mengidentifikasi kehamilan berisiko tinggi. Deteksi dini terhadap kondisi janin yang abnormal memungkinkan intervensi medis yang cepat dan tepat waktu untuk mencegah hasil neonatal yang parah.

Mengacu pada penelitian Shobana & Umamaheswari, (2021) dengan judul “*Prediction of Liver Disease using Gradient Boost Machine Learning Techniques with Feature Scaling*”, ada beberapa algoritma yang diterapkan yaitu *Gradient Boost*, *XGBoost*, *LGBM Classifier*, *CatBoost*, *Logistic Regression*, *Naive Bayes*, *Support Vector Machine*, *K-Nearest Neighbors*, *Decision Tree*, *Random Forest*, dan *Multilayer Perceptron* yang diterapkan dalam *dataset* penyakit liver. Hasil dari penelitian menunjukkan bahwa algoritma *Gradient Boosting Machine* memiliki akurasi tertinggi yaitu 94%. Selain itu, pada penelitian tersebut

disarankan untuk melakukan pengujian pada kumpulan data yang lebih besar untuk mencapai akurasi yang lebih tinggi. Dalam konteks ini, data rekam medis RSUD Al Ihsan Provinsi Jawa Barat memiliki keunggulan dengan memiliki data yang beragam dengan cakupan data yang banyak, memungkinkan penelitian ini untuk bekerja dengan sampel data yang representatif dan signifikan. Dengan kumpulan data yang beragam, penelitian ini dapat memberikan wawasan yang lebih mendalam tentang hubungan antara variabel-variabel terutama yang berkaitan dengan penyakit diabetes, serta menghasilkan model prediksi yang lebih akurat dan bermanfaat bagi pasien dan praktisi medis. Oleh sebab itu, penelitian ini berfokus pada penggunaan algoritma *Gradient Boosting Machine* pada data rekam medis RSUD Al Ihsan Provinsi Jawa Barat dengan dengan jenis penyakit diabetes melitus.

I.2 Rumusan Masalah

Rumusan masalah yang mendasari penelitian ini adalah:

1. Bagaimana hasil penerapan *machine learning* untuk prediksi penyakit diabetes melitus menggunakan algoritma *Gradient Boosting Machine* pada data rekam medis RSUD Al Ihsan Provinsi Jawa Barat?
2. Apa faktor-faktor yang paling mempengaruhi penyakit diabetes melitus yang dialami oleh pasien di RSUD Al Ihsan Provinsi Jawa Barat berdasarkan penerapan algoritma *Gradient Boosting Machine*?

I.3 Tujuan Penelitian

Berdasarkan rumusan masalah, penelitian ini bertujuan untuk:

1. Mengetahui hasil penerapan *machine learning* untuk prediksi penyakit diabetes melitus menggunakan algoritma *Gradient Boosting Machine* pada data rekam medis RSUD Al Ihsan Provinsi Jawa Barat.
2. Mengetahui faktor-faktor yang paling mempengaruhi penyakit diabetes melitus yang dialami oleh pasien di RSUD Al Ihsan Provinsi Jawa Barat berdasarkan penerapan algoritma *Gradient Boosting Machine*.

I.4 Manfaat Penelitian

Manfaat penelitian ini yaitu:

1. Bagi masyarakat luas, hasil penelitian ini digunakan untuk diagnosis dini penyakit diabetes melitus yang dapat meningkatkan kesadaran masyarakat akan pentingnya pencegahan dan deteksi dini. Selain itu, informasi mengenai faktor-faktor yang paling mempengaruhi diabetes melitus dapat membantu masyarakat untuk mengadopsi gaya hidup yang lebih sehat.
2. Bagi RSUD Al Ihsan Provinsi Jawa Barat, penelitian ini dapat menjadi acuan atau referensi dalam peringatan dini untuk meningkatkan diagnosis penyakit diabetes melitus yaitu dengan menerapkan *machine learning*. Selain itu, dapat mengetahui faktor yang paling mempengaruhi penyakit diabetes.
3. Bagi Universitas Telkom, hasil penelitian ini dapat berkontribusi pada pengetahuan ilmiah terutama dalam bidang *machine learning*, *data mining*, atau secara garis besar pada bidang kesehatan. Selain itu, dengan penelitian ini dapat membuka peluang kolaborasi lebih lanjut antara universitas dengan pihak rumah sakit.
4. Bagi peneliti lain, hasil penelitian ini dapat menjadi referensi dalam literatur ilmiah dan juga memperluas atau mengembangkan teori dalam bidang yang relevan. Selain itu, penelitian ini dapat menjadi landasan bagi peneliti lain untuk melanjutkan studi lebih lanjut, menggali aspek-aspek lebih mendalam dalam penggunaan *machine learning* pada data rekam medis.

I.5 Batasan Penelitian

Ruang lingkup yang menjadi batasan penelitian ini yaitu data rekam medis RSUD Al Ihsan Provinsi Jawa Barat dari tahun 2020 sampai dengan 2023. Batasan ini memastikan bahwa data yang digunakan adalah terkini dan relevan, mencakup periode yang cukup untuk analisis tren dan pola dalam diagnosis diabetes melitus.

I.6 Sistematika Penelitian

Penelitian ini diuraikan dengan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Bab I memuat latar belakang dari topik penelitian, rumusan masalah, tujuan penelitian, batasan penelitian, serta manfaat dari penelitian. Kemudian pada akhir bab ini dibahas mengenai sistematika penulisan yang menjadi acuan penulis dalam penyusunan isi dari karya ilmiah ini.

BAB II TINJAUAN PUSTAKA

Bab II berisi literatur atau landasan teori yang relevan dengan topik pembahasan pada penelitian. Landasan teori secara berurut menjelaskan tentang penyakit diabetes, data rekam medis, machine learning, algoritma Gradient Boosting Machine dan teknik evaluasi yang digunakan. Selain dibahas juga penelitian terdahulu yang berkaitan dengan penelitian terkait.

BAB III METODOLOGI PENELITIAN

Bab III membahas terkait pengembangan model konseptual, metode penelitian, dan sistematika penelitian. Pembahasan dilakukan secara berurut dari konsep yang digunakan berkaitan dengan penelitian yang dilakukan hingga alur sistematikan penyelesaian penelitian.

BAB IV ANALISIS DAN PERANCANGAN

Bab IV membahas terkait pokok bahasan tahapan analisis dan perancangan. Tahapan analisis membahas langkah-langkah yang berkaitan dengan pengumpulan hingga pemrosesan data yang digunakan, serta tahapan perancangan yang membahas tentang proses membangun model klasifikasi dengan menggunakan

algoritma *Gradient Boosting Machine*, hingga tahapan evaluasi model.

BAB V IMPLEMENTASI DAN PENGUJIAN

Bab V berisi proses klasifikasi dengan penerapan algoritma *Gradient Boosting Machine*. Hasil pengujian meliputi beberapa metode yaitu *splitting data*, *grid search* dan *k-fold cross validation*. Selain itu, terdapat evaluasi dari penggunaan algoritma yang yang dipakai, meliputi *confusion matrix* dan kurva ROC (AUC).

BAB VI KESIMPULAN DAN SARAN

Bab VI dipaparkan terkait kesimpulan dari hasil yang telah diperoleh dari proses eksplorasi pengujian pada penelitian ini dan diberikan juga saran serta rekomendasi untuk penelitian yang akan datang. Kesimpulan dan saran merujuk kepada hasil analisis keseluruhan yang telah dilakukan dan diuraikan pada bab-bab sebelumnya.