

### Slide Opening

Assalamualaikum warahmatullahi  
wabarakatuh

yang saya hormati

Ibu YANI PARTI ASTUTI, S.Si, M.Kom

Selaku Ketua Dosen Penguji

Bapak JUNTA ZENIARJA, M.Kom

Selaku Dosen Penguji 1

Bapak HARUN AL AZIES, S.Stat., M.Stat.

Selaku Dosen Penguji 2

Dan Bapak ABU SALAM, M.Kom

Selaku Dosen Pembimbing saya

Terimakasih atas kesempatan yang diberikan  
kepada saya untuk melaksanakan sidang tugas  
akhir 2 pada pagi hari ini.

Sebelumnya perkenalkan nama saya Yohanes  
Dimas Pratama dengan NIM A11.2021.13254.

Saya akan mempresentasikan mengenai hasil  
penelitian dan hasil publikasi saya yang  
berjudul Comparison of Data Normalization  
Techniques on KNN Classification Performance  
for Pima Indians Diabetes Dataset.

### Slide Jurnal

Berikut adalah informasi mengenai jurnal saya yang telah di publikasikan di JAIC yaitu Journal of Applied Informatics and Computing (JAIC) yang sudah terakreditasi Sinta 3 pada Volume 9 Number 3 bulan Juni 2025.

Artikel jurnal ini dapat diakses secara online melalui link atau melakukan scan pada QR Code tersebut.

### Slide Penelitian Terkait

Berikut adalah beberapa penelitian terdahulu yang mendukung latar belakang saya dalam mengambil metode dan menjadi acuan bagi penelitian saya.

Pada tahap preprocessing ini saya akan melakukan berbagai pemeriksaan yaitu pemeriksaan tipe data, missing values, dan data duplikat. Pemeriksaan tipe data ini bertujuan untuk memastikan agar tipe data yang berada di setiap fitur ini merupakan tipe data

### Slide Evaluasi

Akan dilakukan pengujian untuk mengevaluasi dan menentukan teknik normalisasi yang paling optimal dalam meningkatkan kinerja model klasifikasi, dengan menganalisis berbagai variasi nilai  $K$  serta dampaknya terhadap hasil yang diperoleh.

Tanpa normalisasi, akurasi maksimal hanya mencapai 0.7597 dan cenderung fluktuatif. Precision, recall, dan F1-score pun rendah, menunjukkan bahwa perhitungan jarak terganggu akibat perbedaan skala antar fitur. Hal ini membuat model sulit mengenali pola secara akurat.

Min-Max Scaling memberikan hasil terbaik di semua metrik, dengan akurasi tertinggi sebesar 0.8117. ROC AUC mencapai 0.8050, memperlihatkan kekuatan model dalam membedakan kelas positif dan kelas negatif secara menyeluruh.

Metrik lainnya seperti precision, recall, F1-score, dan specificity juga menunjukkan peningkatan signifikan. Ini karena skala semua fitur disamakan ke rentang 0–1, sehingga perhitungan jarak menjadi lebih proporsional.

Z-Score Scaling juga menunjukkan hasil yang kompetitif, dengan akurasi mencapai 0.7922. Meskipun beberapa metrik evaluasi masih sedikit di bawah Min-Max Scaling.

Teknik ini menstandarisasi data berdasarkan rata-rata dan standar deviasi, cocok untuk data berdistribusi normal dan mampu menjaga keseimbangan antar fitur.

Sebaliknya, Decimal Scaling justru menghasilkan performa terendah bahkan lebih rendah dari model yang tidak dilakukan normalisasi. Akurasi hanya 0.7403 dan metrik lainnya stagnan.

Karena hanya menggeser titik desimal tanpa menyesuaikan distribusi data, fitur tetap memiliki skala yang tidak seimbang, menyebabkan perhitungan jarak menjadi tidak akurat.

### Slide Uji Statistik

Hasil uji statistik menunjukkan bahwa Min-Max Scaling secara konsisten memberikan peningkatan signifikan pada semua metrik evaluasi, dengan P-Value di bawah 0,05. Peningkatan ini terjadi pada Accuracy, Precision, Recall, F1-Score, dan Specificity, yang menunjukkan bahwa metode ini meningkatkan performa model KNN secara valid dan stabil.

Z-Score Scaling tidak menunjukkan signifikansi pada metrik apapun, kemungkinan karena variabilitas hasil atau ukuran sampel yang terbatas.

Sedangkan Decimal Scaling memperlihatkan signifikansi pada beberapa metrik seperti Precision, Recall, dan F1-Score, meskipun performanya tidak merata dan lebih rendah dibanding Min-Max Scaling.

Secara keseluruhan, Min-Max Scaling terbukti sebagai teknik normalisasi yang paling unggul dan direkomendasikan untuk meningkatkan performa model KNN pada dataset Pima Indians Diabetes.

### Mengapa normalisasi data penting dalam klasifikasi KNN?

Normalisasi data sangat penting dalam KNN karena algoritma ini menghitung kedekatan antar data menggunakan perhitungan jarak, seperti Euclidean Distance. Jika data memiliki skala atau satuan yang berbeda, hal ini dapat menyebabkan ketidakseimbangan dalam perhitungan jarak dan menurunkan kinerja model. Oleh karena itu, normalisasi dilakukan untuk memastikan bahwa setiap fitur berada dalam skala yang seragam, sehingga perhitungan jarak antar data menjadi lebih akurat.

### Kenapa pilih metode normalisasi tersebut?

Metode normalisasi yang dipilih, yaitu Min-Max Scaling, Z-Score Scaling, dan Decimal Scaling, dipilih karena ketiganya relevan dengan perhitungan Euclidean Distance yang digunakan dalam algoritma KNN.

### Kenapa relevan dengan perhitungan Euclidean Distance?

Metode normalisasi tersebut relevan dengan perhitungan Euclidean Distance karena teknik normalisasi ini menyamakan skala antar fitur. Dalam KNN, perhitungan jarak antar data menggunakan Euclidean Distance, yang mengukur kedekatan titik data. Jika fitur memiliki skala yang berbeda, fitur dengan rentang nilai yang lebih besar akan mendominasi perhitungan jarak, mengurangi akurasi model. Normalisasi memastikan setiap fitur berkontribusi secara seimbang dalam perhitungan jarak, sehingga meningkatkan kinerja model.

### Apa alasan khusus Anda memilih dataset Pima Indians Diabetes untuk penelitian ini dibanding dengan dataset diabetes lain?

Saya memilih dataset Pima Indians Diabetes karena sudah banyak digunakan dalam penelitian sebelumnya, memiliki fitur medis yang relevan untuk klasifikasi diabetes, dan ukurannya yang relatif sederhana dengan hanya 768 data dan 8 fitur, sehingga mudah untuk dikelola dan digunakan dalam eksperimen. Ukuran dataset yang tidak terlalu besar memungkinkan untuk eksplorasi teknik normalisasi dan pemodelan dengan lebih cepat.

### Kenapa pilih dataset dengan penyakit diabetes?

Saya memilih dataset Pima Indians Diabetes karena diabetes adalah penyakit yang prevalensinya semakin meningkat dan memerlukan deteksi dini. Dataset ini relevan untuk mengklasifikasikan risiko diabetes berdasarkan fitur medis seperti kadar glukosa dan BMI, serta sudah banyak digunakan dalam penelitian machine learning, termasuk penelitian ini dengan algoritma K-Nearest Neighbors (KNN).

### Kenapa melakukan penelitian ini? Apa bedanya dengan penelitian sebelumnya / penelitian terkait?

Penelitian ini dilakukan untuk membandingkan teknik normalisasi data terhadap kinerja model KNN dalam klasifikasi diabetes. Perbedaannya dengan penelitian sebelumnya adalah fokus pada tiga teknik normalisasi yang relevan dengan Euclidean Distance dan evaluasi menggunakan berbagai metrik serta uji statistik untuk memastikan signifikansi perbedaannya.

Di penelitian terkait ada penelitian normalisasi dengan algoritma SVM dan ANN, kenapa lebih pilih KNN?

- Untuk dataset ini yang berukuran kecil hingga menengah, KNN memberikan hasil yang sangat baik, dengan normalisasi meningkatkan efektivitasnya karena data yang lebih seragam menghasilkan prediksi yang lebih akurat. Karena berdasarkan literasi yang saya pelajari, SVM dan ANN lebih cocok untuk dataset besar yang membutuhkan banyak data untuk mengoptimalkan model.
- KNN adalah algoritma yang relatif sederhana dan efektif untuk masalah klasifikasi yang tidak memerlukan banyak parameter tuning. Berbeda dengan SVM dan ANN yang memerlukan pengaturan parameter lebih rumit, KNN hanya menghitung jarak antara data uji dan data latih. Maka dari itu normalisasi data sangat penting dalam KNN untuk memastikan perhitungan jarak antar data dapat dilakukan dengan akurat, terutama pada dataset yang beragam.

Akurasi tertinggi min-max scaling saya memang lebih baik dari penelitian terkait dengan dataset dan metode yang sama, tetapi kenapa akurasi stuck di 80? Bisa naik lagi gak?

Akurasi yang stuck di sekitar 81% bisa disebabkan oleh beberapa faktor, seperti pemilihan parameter K yang kurang optimal, keterbatasan model KNN dalam menangani kompleksitas data, atau mungkin fitur yang digunakan masih belum sepenuhnya optimal. Untuk meningkatkan akurasi, saya bisa mencoba mencari nilai K yang lebih tepat menggunakan teknik cross-validation, serta mengeksplorasi algoritma lain yang mungkin lebih efektif, seperti Random Forest atau SVM. Selain itu, pengolahan data lebih lanjut atau penambahan data baru juga bisa membantu meningkatkan performa model.

Apakah ada batasan, kelemahan, dan saran untuk penelitian selanjutnya?

- Batasan dalam penelitian ini termasuk fokus pada tiga teknik normalisasi saja, sehingga ada kemungkinan teknik lain yang lebih efektif belum dieksplorasi. Selain itu, dataset yang digunakan hanya mencakup perempuan dari etnis Pima, yang mungkin tidak sepenuhnya representatif untuk populasi secara umum. Terakhir, KNN sebagai algoritma berbasis jarak mungkin kurang efektif untuk data dengan dimensi tinggi atau distribusi fitur yang sangat berbeda.
- Terdapat beberapa kelemahan dalam penelitian ini, di antaranya dataset yang terbatas dan tidak seimbang antar kelas, yang dapat mempengaruhi hasil klasifikasi. Selain itu, pemilihan nilai K yang terbatas juga menjadi faktor yang membatasi akurasi.
- Untuk penelitian selanjutnya, saran yang dapat diberikan adalah mencoba teknik normalisasi lain seperti Robust Scaling, serta menguji algoritma lain seperti SVM atau Random Forest yang mungkin lebih efektif dalam menangani kompleksitas data. Penambahan data atau penggunaan cross-validation untuk memilih nilai K yang optimal juga dapat meningkatkan performa model.

### Fitur apa yang paling penting? Lalu kenapa fitur itu memiliki kontribusi yang paling penting?

- Fitur yang paling penting dalam penelitian ini adalah Glucose, BMI, Age, dan DiabetesPedigreeFunction. Fitur-fitur ini memiliki kontribusi terbesar dalam klasifikasi diabetes karena mereka berhubungan langsung dengan risiko diabetes secara medis. Misalnya, kadar glukosa dalam darah adalah indikator utama dalam mendiagnosis diabetes, BMI menunjukkan kondisi obesitas yang meningkatkan risiko diabetes, dan Age serta DiabetesPedigreeFunction menggambarkan faktor usia dan riwayat keluarga, yang keduanya berperan dalam peningkatan risiko penyakit ini.
- Dari sisi rentang nilai dalam dataset, fitur-fitur ini juga menunjukkan variasi yang cukup besar, yang memungkinkan model untuk membedakan antara individu yang berisiko tinggi dan rendah. Misalnya, glukosa memiliki rentang nilai yang luas, yang memberikan informasi lebih banyak tentang status kesehatan pasien. Oleh karena itu, meskipun keduanya, sisi medis dan rentang nilai dalam dataset, berkontribusi, faktor medis lebih mendalam dalam menentukan pentingnya fitur tersebut.

### Fitur apa aja yang dihapus? Kenapa memutuskan untuk menghapus fitur itu?

- Keputusan untuk menghapus fitur-fitur ini diambil berdasarkan hasil pemilihan fitur menggunakan metode Random Forest. Fitur-fitur ini memiliki nilai penting yang rendah (feature importance) dan tidak memberikan kontribusi signifikan terhadap model. Penghapusan fitur yang tidak relevan ini bertujuan untuk meningkatkan efisiensi model, mengurangi risiko overfitting, dan memastikan bahwa model hanya berfokus pada fitur yang lebih relevan untuk klasifikasi diabetes, seperti Glucose, BMI, Age, dan DiabetesPedigreeFunction.
- Dari sisi medis, fitur seperti SkinThickness, Insulin, Pregnancies, dan BloodPressure dihapus karena meskipun terkait dengan diabetes, mereka tidak memiliki kontribusi yang kuat dalam mendeteksi atau memprediksi diabetes pada dataset ini. Fitur-fitur ini mungkin tidak cukup relevan dalam memberikan gambaran menyeluruh tentang kondisi kesehatan pasien yang berkaitan langsung dengan diabetes, dibandingkan dengan fitur seperti Glucose dan BMI yang lebih berhubungan langsung dengan risiko diabetes.
- Dari sisi rentang nilai, fitur-fitur yang dihapus memiliki rentang nilai yang terbatas atau variabilitas yang rendah, yang membuatnya kurang berguna dalam perhitungan model. Fitur seperti SkinThickness dan BloodPressure memiliki variasi nilai yang tidak cukup besar, sehingga tidak memberikan informasi yang cukup untuk membedakan antara individu yang berisiko tinggi atau rendah terkena diabetes. Oleh karena itu, fitur-fitur ini tidak memberi dampak signifikan pada hasil klasifikasi.

### Dataset isinya diabetes cewe, bagaimana kalau diterapkan di cowo? Bagaimana pengaruhnya?

Jika diterapkan pada data pria, hasilnya bisa berbeda karena perbedaan biologis, seperti metabolisme dan kadar glukosa, yang memengaruhi faktor risiko diabetes. Model yang dilatih dengan data perempuan mungkin kurang akurat untuk pria, sehingga perlu penyesuaian atau pelatihan model dengan data yang mencakup kedua jenis kelamin untuk meningkatkan generalisasi dan akurasi.

Kenapa uji statistik Z-Score berbanding terbalik dengan evaluasi. Di evaluasi dia kompetitif secara metrik evaluasi dengan min-max scaling, tetapi kenapa di uji statistik itu malah tidak ada yang signifikan di setiap metrik evaluasi?

Z-Score Scaling menunjukkan hasil yang kompetitif dalam evaluasi metrik, tetapi uji statistiknya tidak signifikan karena **variabilitas hasil** atau **ukuran sampel yang terbatas**. Variabilitas ini bisa terjadi karena distribusi data yang tidak konsisten atau ukuran sampel yang terlalu kecil, sehingga perbedaan yang ada tidak cukup besar untuk terdeteksi sebagai signifikan secara statistik. Hal ini mengindikasikan bahwa meskipun Z-Score Scaling memberikan performa yang baik dalam beberapa metrik, perbedaannya tidak cukup stabil atau konsisten untuk dianggap signifikan dalam uji statistik.

Apakah ada uji statistik lain yang dapat digunakan untuk memperkuat temuan Anda?

Untuk memperkuat temuan ini, uji statistik lain yang dapat digunakan adalah ANOVA (Analysis of Variance) jika lebih dari dua teknik normalisasi ingin dibandingkan.

Mengapa Anda memilih Paired T-Test sebagai uji statistik untuk membandingkan teknik normalisasi? Apakah Anda mempertimbangkan menggunakan uji statistik non-parametrik seperti Wilcoxon Signed-Rank Test, terutama jika data Anda tidak mengikuti distribusi normal?

Paired T-Test dipilih karena saya membandingkan dua kondisi yang berhubungan langsung (normalisasi vs tanpa normalisasi) pada data yang berpasangan. Namun, jika data tidak memenuhi asumsi normalitas, **Wilcoxon Signed-Rank Test** bisa menjadi alternatif yang lebih tepat karena tidak bergantung pada asumsi distribusi normal dan lebih cocok untuk data ordinal atau data yang tidak terdistribusi normal.

Jika Anda ingin meningkatkan kekuatan statistik dalam penelitian ini, apakah ada perubahan dalam desain eksperimen atau metode analisis yang akan Anda lakukan?

Untuk meningkatkan kekuatan statistik, saya akan **menambah ukuran sampel** agar hasil uji statistik lebih robust dan dapat mendeteksi perbedaan kecil dengan lebih jelas.

Mengapa Anda menggunakan beberapa metrik evaluasi seperti akurasi, precision, recall, F1-score, specificity, dan ROC AUC dalam penelitian ini?

Berbagai metrik evaluasi digunakan untuk memberikan gambaran yang lebih komprehensif tentang kinerja model. Akurasi mengukur seberapa baik model mengklasifikasikan data secara keseluruhan. Precision dan recall memberikan gambaran lebih mendalam tentang kemampuan model dalam mengklasifikasikan kelas positif dan negatif dengan benar. F1-score menggabungkan precision dan recall untuk memberikan keseimbangan, sedangkan specificity mengukur kemampuan model dalam mengidentifikasi data negatif dengan benar. ROC AUC mengukur kemampuan model dalam membedakan antara kedua kelas secara keseluruhan.

Mengapa Anda berpikir Decimal Scaling menunjukkan hasil yang buruk, dan apa yang dapat dilakukan untuk meningkatkan kinerjanya di penelitian mendatang?

Decimal Scaling menunjukkan hasil yang buruk karena metode ini mengubah rentang nilai fitur dengan membaginya dengan pangkat sepuluh. Hal ini tidak memperhitungkan distribusi data atau penyebaran nilai, yang mengakibatkan ketidakseimbangan dalam perhitungan jarak. Akibatnya, perhitungan jarak dalam KNN menjadi bias dan tidak efektif. Untuk meningkatkan kinerjanya, kita bisa mencoba teknik normalisasi lainnya yang lebih memperhitungkan distribusi dan variasi data, atau menggunakan metode yang lebih canggih seperti Robust Scaling.

### Bagaimana implementasi model ini dalam dunia medis?

Model ini dapat digunakan dalam dunia medis untuk **deteksi dini diabetes**, membantu **tenaga medis dalam pengambilan keputusan**, **memantau kondisi pasien**, dan **program pencegahan** dengan mengidentifikasi risiko diabetes berdasarkan data medis seperti kadar glukosa, BMI, dan usia. Implementasinya dapat mempercepat diagnosis dan memberikan rekomendasi pengelolaan lebih awal.

### Implementasi programnya?

Implementasi program ini dalam dunia medis melibatkan pengumpulan data pasien, pra-pemrosesan data, penerapan model KNN untuk klasifikasi risiko diabetes, dan memberikan rekomendasi tindak lanjut berdasarkan hasil klasifikasi. Program ini dapat diintegrasikan dengan sistem manajemen rumah sakit untuk mempermudah deteksi dini dan pengambilan keputusan medis.

### Jika Anda harus mengembangkan penelitian ini lebih lanjut, faktor atau dataset tambahan apa yang akan Anda masukkan?

Untuk mengembangkan penelitian ini, saya akan mempertimbangkan untuk memasukkan dataset yang lebih beragam dengan lebih banyak variabel medis yang relevan, seperti kadar kolesterol, riwayat keluarga, atau faktor gaya hidup. Menambahkan data dari populasi yang lebih besar dan lebih beragam akan membantu memperbaiki generalisasi model dan meningkatkan akurasi prediksi diabetes.

### Untuk parameter $k$ , yaitu 1, 3, 5, 7, 9, dan 11, yang dipilih secara manual ya, kenapa memilih parameter itu?

Parameter  $K$  (1, 3, 5, 7, 9, dan 11) dipilih untuk menguji keseimbangan antara **overfitting** dan **underfitting**. Nilai  $K$  yang lebih kecil cenderung menyebabkan overfitting, sedangkan nilai  $K$  yang lebih besar dapat menyebabkan underfitting. Pemilihan nilai-nilai ini untuk mencari  $K$  yang memberikan kinerja optimal pada model.

### Kenapa pembagian data menggunakan proporsi 80% data latih dan 20% data uji?

Pembagian data dengan proporsi **80% data latih** dan **20% data uji** dipilih untuk memberikan cukup data untuk melatih model (80%) sementara masih menyisakan sebagian data (20%) untuk menguji dan mengevaluasi kinerja model. Proporsi ini umum digunakan karena memungkinkan model untuk belajar dari banyak data, sambil tetap memiliki data uji yang cukup untuk menguji generalisasi model terhadap data yang belum pernah dilihat sebelumnya.