Stemming Results

# Processing Steps:

## Tokenization:

Tokens: penerapan, metode, k, nearest, neighbors, k, nn, untuk, klasifikasi, risiko, diabetes, pada, dataset, kesehatan, disusun, oleh, emelsha, viadra, naufal, zaidan, prodi, informatika, fakultas, teknologi, industri, institut, teknologi, nasional, bandung, k, nearest, neighbors, k, nn, adalah, algoritma, supervised, learning, yang, banyak, digunakan, dalam, klasifikasi, dan, regresi, algoritma, ini, bekerja, dengan, memanfaatkan, kedekatan, atau, kemiripan, antara, data, baru, dengan, data, yang, sudah, ada, untuk, menentukan, kelas, atau, nilai, prediksinya, dalam, penerapannya, k, nn, dimulai, dengan, menetapkan, parameter, k, yang, merupakan, jumlah, tetangga, terdekat, yang, akan, dipertimbangkan, saat, mengklasifikasikan, data, baru, pemilihan, nilai, k, ini, sangat, penting, karena, k, yang, terlalu, kecil, bisa, membuat, model, terlalu, peka, terhadap, data, tertentu, overfitting, sementara, k, yang, terlalu, besar, bisa, membuat, model, kehilangan, sensitivitas, terhadap, pola, lokal, pada, data, proses, utama, k, nn, melibatkan, perhitungan, jarak, antara, data, baru, dan, setiap, data, dalam, dataset, yang, sudah, ada, jarak, ini, biasanya, dihitung, dengan, menggunakan, metrik, seperti, euclidean, distance, manhattan, distance, atau, minkowski, distance, euclidean, distance, adalah, salah, satu, metode, yang, paling, umum, digunakan, di, mana, jarak, diukur, secara, lurus, antara, dua, titik, dalam, ruang, fitur, untuk, dataset, yang, memiliki, fitur, dengan, dimensi, yang, berbeda, beda, normalisasi, data, sering, kali, diperlukan, untuk, memastikan, jarak, t, ersebut, relevan, setelah, semua, jarak, dihitung, algoritma, kemudian, memilih, k, data, terdekat, yaitu, data, yang, memiliki, jarak, terkecil, terhadap, data, baru, pada, tahap, akhir, algoritma, menentukan, kelas, data, baru, berdasarkan, mayoritas, kelas, dari, k, tetangga, terdekat, tersebut, misa, lnya, jika, mayoritas, dari, k, tetangga, berada, dalam, kelas, diabetes, positif, maka, data, baru, juga, akan, diklasifikasikan, sebagai, diabetes, positif, kelebihan, utama, dari, k, nn, adalah, kesederhanaannya, karena, tidak, memerlukan, proses, pelatihan, yang, rumit, namun, kel, emahannya, adalah, algoritma, ini, bisa, menjadi, lambat, jika, ukuran, dataset, sangat, besar, karena, harus, menghitung, jarak, antara, data, baru, dengan, setiap, data, dalam, dataset, sejarah, algoritma, k, nearest, neighbors, k, nn, berakar, pada, perkembangan, awal, kecerdasan, buatan, dan, analisis, data, statistik, meskipun, konsep, dasar, dari, algoritma, ini, telah, dikenal, sejak, awal, abad, ke, perkembangan, signifikan, dalam, teori, klasifikasi, dan, pengenalan, pola, baru, terjadi, pada, pertengahan, hingga, akhir, an, berikut, adalah, beberapa, poin, penting, dalam, sejarah, algoritma, k, nn, penemuan, oleh, evelyn, fix, dan, joseph, hodges, •, algoritma, k, nn, pertama, kali, diperkenalkan, oleh, evelyn, fix, dan, joseph, hodges, pada, tahun, dalam, makalah, mereka, berjudul, discriminatory, analysis, nonparametric, discrimination, consistency, properties, •, mereka, bekerja, di, universitas, california, berkeley, dan, fokus, utama, penelitian, mereka, adalah, pengembangan, metode, non, parametrik, untuk, klasifikasi, •, makalah, ini, membahas, penggunaan, jarak, sebagai, dasar, untuk, mengklasifikasikan, data, baru, dengan, melihat, tetangga, terdekat, dalam, kumpulan, data, yang, telah, diberi, label, penerapan, dalam, pengenalan, pola, •, pada, tahun, algoritma, k, nn, semakin, populer, ketika, digunakan, dalam, buku, berjudul, the, theory, of, pattern, recognition, oleh, thomas, m, cover, dan, peter, e, hart, •, cover, dan, hart, mengembangkan, teori, dasar, dari, k, nn, dan, menunjukkan, bahwa, metode, ini, bisa, sangat, efektif, untuk, pengenalan, pola, tanpa, memerlukan, asumsi, yang, rumit, mengenai, distribusi, data, •, penelitian, mereka, menginspirasi, banyak, ilmuwan, lain, untuk, mengembangkan, metode, klasifikasi, berbasis, tetangga, terdekat, an, perkembangan, di, bidang, pengklasifikasian, •, algoritma, k, nn, mulai, banyak, digunakan, dalam, berbagai, aplikasi, pengklasifikasian, terutama, dalam, bidang, statistik, dan, pembelajaran, mesin, •, k, nn, diadopsi, sebagai, pendekatan, sederhana, namun, kuat, untuk, klasifikasi, data, aplikasi, utamanya, adalah, di, bidang, bioinformatika, pengenalan, wajah, dan, pengenalan, tulisan, tangan, an, penggunaan, dalam, komputasi, •, pada, era, ini, k, nn, semakin, populer, dengan, munculnya, komputasi, yang, lebih, cepat, dan, lebih, terjangkau, •, penggunaan, k, nn, meluas, ke, bidang, yang, lebih, kompleks, seperti, komputer, vision, pengenalan, suara, dan, diagnosis, medis, an, kebangkitan, bersama, pembelajaran, mesin, •, dengan, berkembangnya, bidang, pembelajaran, mesin, machine, learning, k, nn, menjadi, salah, satu, algoritma, dasar, yang, sering, diajarkan, dan, digunakan, dalam, pemodelan, awal, data, •, k, nn, banyak, digunakan, untuk, klasifikasi, data, di, berbagai, industri, seperti, keuangan, kesehatan, dan, perdagangan, elektronik, hingga, saat, ini, •, k, nn, dianggap, sebagai, algoritma, yang, mudah, dipahami, dan, diterapkan, yang, tetap, relevan, hingga, kini, dalam, berbagai, aplikasi, pembelajaran, mesin, •, meskipun, algoritma, ini, sederhana, dan, memiliki, beberapa, keterbatasan, seperti, rentan, terhadap, overfitting, dan, performa, yang, lambat, pada, dataset, yang, besar, k, nn, masih, populer, sebagai, metode, awal, dalam, eksplorasi, data, dalam, algoritma, k, nearest, neighbors, k, nn, salah, satu, metrik, yang, umum, digunakan, untuk, mengukur, jarak, antar, data, adalah, euclidean, distance, metode, ini, menghitung, jarak, lurus, antara, dua, titik, dalam, ruang, fitur, dan, sering, digunakan, karena, kesederhanaannya, dalam, mengukur, kedekatan, antar, titik, rumus, euclidean, distance, antara, dua, titik, x, dan, y, dengan, fitur, x1, x2, xn, dan, y1, y2, yn, adalah, penjelasan, •, x, dan, y, adalah, dua, data, atau, titik, yang, jaraknya, akan, dihitung, •, xi, dan, yi, adalah, nilai, fitur, ke, i, untuk, titik, x, dan, y, •, n, adalah, jumlah, fitur, yang, ada, dalam, dataset, contoh, kasus, diabetes, merupakan, salah, satu, penyakit, yang, banyak, dialami, oleh, masyarakat, dan, deteksi, dini, sangat, penting, untuk, mencegah, komplikasi, lebih, lanjut, dalam, studi, ini, kita, akan, menggunakan, algoritma, k, nearest, neighbors, k, nn, untuk, mengklasifikasikan, apakah, seseorang, berisiko, terkena, diabetes, berdasarkan, beberapa, faktor, kesehatan, tabel, data, sampel, data, sampel, no, gula, darah, insulin, bmi, usia, diabet, tidak, ya, ya, ya, ya, ya, ya, tidak, ya, tidak, ya, ya, tidak, tidak, ya, tidak, tidak, ya, tidak, ya, tidak, tidak, tidak, ya, tidak, tidak, tidak, penyelesaian, dari, data, tersebut, k, ita, ambil, data, untuk, menjadi, acuan, pengujian, terhadap, seluruh, data, tabel, data, uji, data, uji, gula, darah, insulin, bmi, usia, diabet, ya, tidak, ya, tabel, hasil, dari, pengujian, data, data, uji, jarak, euclidian, urutan, diabet, tidak, ya, ya, ya, tidak, tabel, hasil, dari, pengujian, data, data, uji, jarak, euclidian, urutan, diabet, ya, tidak, tidak, tidak, tidak, tabel, hasil, dari, pengujian, data, data, uji, jarak, euclidian, urutan, diabet, ya, ya, ya, ya, tidak, dalam, pengujian, ini, kita, menggunakan, algoritma, k, nearest, neighbors, k, nn, dengan, k, untuk, menentukan, risiko, diabetes, pada, tiga, data, uji, berdasarkan, jarak, euclidean, terhadap, data, sampel, data, uji, lima, tetangga, terdekat, untuk, data, uji, pertama, dengan, hasil, klasifikasi, adalah, •, jarak, dari, data, sampel, no, ya, •, jarak, dari, data, sampel, no, tidak, •, jarak, dari, data, sampel, no, ya, •, jarak, dari, data, sampel, no, ya, •, jarak, dari, data, sampel, no, tidak, dari, lima, tetangga, terdekat, tiga, di, antaranya, memiliki, hasil, ya, dan, dua, memiliki, hasil, tidak, karena, mayoritas, hasil, adalah, ya, data, uji, pertama, diklasifikasikan, sebagai, berisiko, diabetes, data, uji, lima, tetangga, terdekat, untuk, data, uji, kedua, dengan, hasil, klasifikasi, adalah, •, jarak, dari, data, sampel, no, tidak, •, jarak, dari, data, sampel, no, tidak, •, jarak, dari, data, sampel, no, tidak, •, jarak, dari, data, sampel, no, tidak, •, jarak, dari, data, sampel, no, ya, dari, lima, tetangga, terdekat, empat, di, antaranya, memiliki, hasil, tidak, dan, satu, memiliki, hasil, ya, dengan, demikian, data, uji, kedua, diklasifikasikan, sebagai, tidak, berisiko, diabetes, data, uji, lima, tetangga, terdekat, untuk, data, uji, ketiga, dengan, hasil, klasifikasi, adalah, •, jarak, dari, data, sampel, no, ya, •, jarak, dari, data, sampel, no, tidak, •, jarak, dari, data, sampel, no, ya, •, jarak, dari, data, sampel, no, ya, •, jarak, dari, data, sampel, no, ya, dari, lima, tetangga, terdekat, empat, memiliki, hasil, ya, dan, satu, memiliki, hasil, tidak, oleh, karena, itu, data, uji, ketiga, diklasifikasikan, sebagai, berisiko, diabetes, ringkasan, dengan, menggunakan, k, hasil, klasifikasi, menunjukkan, bahwa, •, data, uji, pertama, dan, ketiga, diklasifikasikan, sebagai, berisiko, diabetes, •, data, uji, kedua, diklasifikasikan, sebagai, tidak, berisiko, diabetes

## Stemming Process:

### Word: ukuran → ukur

* Step 1: Checking 'ukuran' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'ukuran'
* Removed -an: 'ukuran' → 'ukur'
* Result: Found 'ukur' in dictionary

### Word: buatan → buat

* Step 1: Checking 'buatan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'buatan'
* Removed -an: 'buatan' → 'buat'
* Result: Found 'buat' in dictionary

### Word: kumpulan → kumpul

* Step 1: Checking 'kumpulan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'kumpulan'
* Removed -an: 'kumpulan' → 'kumpul'
* Result: Found 'kumpul' in dictionary

### Word: tulisan → tulis

* Step 1: Checking 'tulisan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'tulisan'
* Removed -an: 'tulisan' → 'tulis'
* Result: Found 'tulis' in dictionary

### Word: tangan → tang

* Step 1: Checking 'tangan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'tangan'
* Removed -an: 'tangan' → 'tang'
* Result: Found 'tang' in dictionary

### Word: xi → x

* Step 1: Checking 'xi' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'xi'
* Removed -i: 'xi' → 'x'
* Result: Found 'x' in dictionary

### Word: yi → y

* Step 1: Checking 'yi' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'yi'
* Removed -i: 'yi' → 'y'
* Result: Found 'y' in dictionary

### Word: acuan → acu

* Step 1: Checking 'acuan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'acuan'
* Removed -an: 'acuan' → 'acu'
* Result: Found 'acu' in dictionary

### Word: urutan → urut

* Step 1: Checking 'urutan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'urutan'
* Removed -an: 'urutan' → 'urut'
* Result: Found 'urut' in dictionary

### Word: urutan → urut

* Step 1: Checking 'urutan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'urutan'
* Removed -an: 'urutan' → 'urut'
* Result: Found 'urut' in dictionary

### Word: urutan → urut

* Step 1: Checking 'urutan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'urutan'
* Removed -an: 'urutan' → 'urut'
* Result: Found 'urut' in dictionary

### Word: ringkasan → ringkas

* Step 1: Checking 'ringkasan' in dictionary
* Step 2: Checking inflection suffixes
* Step 3: Checking derivation suffixes for 'ringkasan'
* Removed -an: 'ringkasan' → 'ringkas'
* Result: Found 'ringkas' in dictionary

# Final Stemmed Text:

penerapan metode k nearest neighbors k nn klasifikasi risiko diabetes dataset kesehatan disusun emelsha viadra naufal zaidan prodi informatika fakultas teknologi industri institut teknologi nasional bandung k nearest neighbors k nn algoritma supervised learning klasifikasi regresi algoritma memanfaatkan kedekatan kemiripan data data menentukan kelas nilai prediksinya penerapannya k nn menetapkan parameter k tetangga terdekat dipertimbangkan mengklasifikasikan data pemilihan nilai k k model peka data overfitting k model kehilangan sensitivitas pola lokal data proses utama k nn melibatkan perhitungan jarak data data dataset jarak dihitung metrik euclidean distance manhattan distance minkowski distance euclidean distance salah metode jarak diukur lurus titik ruang fitur dataset memiliki fitur dimensi berbeda beda normalisasi data kali jarak t ersebut relevan jarak dihitung algoritma memilih k data terdekat data memiliki jarak terkecil data tahap algoritma menentukan kelas data berdasarkan mayoritas kelas k tetangga terdekat misa lnya mayoritas k tetangga kelas diabetes positif data diklasifikasikan diabetes positif kelebihan utama k nn kesederhanaannya proses pelatihan rumit kel emahannya algoritma lambat ukur dataset menghitung jarak data data dataset sejarah algoritma k nearest neighbors k nn berakar perkembangan kecerdasan buat analisis data statistik konsep dasar algoritma dikenal abad perkembangan signifikan teori klasifikasi pengenalan pola pertengahan an poin sejarah algoritma k nn penemuan evelyn fix joseph hodges • algoritma k nn kali diperkenalkan evelyn fix joseph hodges makalah berjudul discriminatory analysis nonparametric discrimination consistency properties • universitas california berkeley fokus utama penelitian pengembangan metode non parametrik klasifikasi • makalah membahas penggunaan jarak dasar mengklasifikasikan data tetangga terdekat kumpul data label penerapan pengenalan pola • algoritma k nn populer buku berjudul the theory of pattern recognition thomas m cover peter e hart • cover hart mengembangkan teori dasar k nn metode efektif pengenalan pola asumsi rumit distribusi data • penelitian menginspirasi ilmuwan mengembangkan metode klasifikasi berbasis tetangga terdekat an perkembangan bidang pengklasifikasian • algoritma k nn aplikasi pengklasifikasian bidang statistik pembelajaran mesin • k nn diadopsi pendekatan sederhana kuat klasifikasi data aplikasi utamanya bidang bioinformatika pengenalan wajah pengenalan tulis tang an penggunaan komputasi • era k nn populer munculnya komputasi cepat terjangkau • penggunaan k nn meluas bidang kompleks komputer vision pengenalan suara diagnosis medis an kebangkitan pembelajaran mesin • berkembangnya bidang pembelajaran mesin machine learning k nn salah algoritma dasar diajarkan pemodelan data • k nn klasifikasi data industri keuangan kesehatan perdagangan elektronik • k nn dianggap algoritma mudah dipahami diterapkan relevan aplikasi pembelajaran mesin • algoritma sederhana memiliki keterbatasan rentan overfitting performa lambat dataset k nn populer metode eksplorasi data algoritma k nearest neighbors k nn salah metrik mengukur jarak data euclidean distance metode menghitung jarak lurus titik ruang fitur kesederhanaannya mengukur kedekatan titik rumus euclidean distance titik x y fitur x1 x2 xn y1 y2 yn penjelasan • x y data titik jaraknya dihitung • x y nilai fitur i titik x y • n fitur dataset contoh diabetes salah penyakit dialami masyarakat deteksi mencegah komplikasi studi algoritma k nearest neighbors k nn mengklasifikasikan berisiko terkena diabetes berdasarkan faktor kesehatan tabel data sampel data sampel no gula darah insulin bmi usia diabet ya ya ya ya ya ya ya ya ya ya ya ya ya penyelesaian data k ita ambil data acu pengujian data tabel data uji data uji gula darah insulin bmi usia diabet ya ya tabel hasil pengujian data data uji jarak euclidian urut diabet ya ya ya tabel hasil pengujian data data uji jarak euclidian urut diabet ya tabel hasil pengujian data data uji jarak euclidian urut diabet ya ya ya ya pengujian algoritma k nearest neighbors k nn k menentukan risiko diabetes data uji berdasarkan jarak euclidean data sampel data uji tetangga terdekat data uji hasil klasifikasi • jarak data sampel no ya • jarak data sampel no • jarak data sampel no ya • jarak data sampel no ya • jarak data sampel no tetangga terdekat memiliki hasil ya memiliki hasil mayoritas hasil ya data uji diklasifikasikan berisiko diabetes data uji tetangga terdekat data uji hasil klasifikasi • jarak data sampel no • jarak data sampel no • jarak data sampel no • jarak data sampel no • jarak data sampel no ya tetangga terdekat memiliki hasil memiliki hasil ya data uji diklasifikasikan berisiko diabetes data uji tetangga terdekat data uji ketiga hasil klasifikasi • jarak data sampel no ya • jarak data sampel no • jarak data sampel no ya • jarak data sampel no ya • jarak data sampel no ya tetangga terdekat memiliki hasil ya memiliki hasil data uji ketiga diklasifikasikan berisiko diabetes ringkas k hasil klasifikasi • data uji ketiga diklasifikasikan berisiko diabetes • data uji diklasifikasikan berisiko diabetes