Stemming Results

# Stemming Process:

## Word: disusun → susun

* Step 1: Checking 'disusun' in dictionary
* Step 2: Checking inflection suffixes for 'disusun'
* Step 3: Checking derivation suffixes for 'disusun'
* Step 4b: Removing prefixes from 'disusun'
* Removed prefix: 'disusun' → 'susun'
* Result: Found 'susun' in dictionary

## Word: terdekat → dekat

* Step 1: Checking 'terdekat' in dictionary
* Step 2: Checking inflection suffixes for 'terdekat'
* Step 3: Checking derivation suffixes for 'terdekat'
* Step 4b: Removing prefixes from 'terdekat'
* Removed prefix: 'terdekat' → 'dekat'
* Result: Found 'dekat' in dictionary

## Word: dihitung → hitung

* Step 1: Checking 'dihitung' in dictionary
* Step 2: Checking inflection suffixes for 'dihitung'
* Step 3: Checking derivation suffixes for 'dihitung'
* Step 4b: Removing prefixes from 'dihitung'
* Removed prefix: 'dihitung' → 'hitung'
* Result: Found 'hitung' in dictionary

## Word: diukur → ukur

* Step 1: Checking 'diukur' in dictionary
* Step 2: Checking inflection suffixes for 'diukur'
* Step 3: Checking derivation suffixes for 'diukur'
* Step 4b: Removing prefixes from 'diukur'
* Removed prefix: 'diukur' → 'ukur'
* Result: Found 'ukur' in dictionary

## Word: terkecil → kecil

* Step 1: Checking 'terkecil' in dictionary
* Step 2: Checking inflection suffixes for 'terkecil'
* Step 3: Checking derivation suffixes for 'terkecil'
* Step 4b: Removing prefixes from 'terkecil'
* Removed prefix: 'terkecil' → 'kecil'
* Result: Found 'kecil' in dictionary

## Word: terdekat → dekat

* Step 1: Checking 'terdekat' in dictionary
* Step 2: Checking inflection suffixes for 'terdekat'
* Step 3: Checking derivation suffixes for 'terdekat'
* Step 4b: Removing prefixes from 'terdekat'
* Removed prefix: 'terdekat' → 'dekat'
* Result: Found 'dekat' in dictionary

## Word: kel → l

* Step 1: Checking 'kel' in dictionary
* Step 2: Checking inflection suffixes for 'kel'
* Step 3: Checking derivation suffixes for 'kel'
* Step 4b: Removing prefixes from 'kel'
* Removed prefix: 'kel' → 'l'
* Result: Found 'l' in dictionary

## Word: ukuran → ukur

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

## Word: berakar → akar

* Step 1: Checking 'berakar' in dictionary
* Step 2: Checking inflection suffixes for 'berakar'
* Step 3: Checking derivation suffixes for 'berakar'
* Step 4b: Removing prefixes from 'berakar'
* Removed prefix: 'berakar' → 'akar'
* Result: Found 'akar' in dictionary

## Word: buatan → buat

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

## Word: dikenal → kenal

* Step 1: Checking 'dikenal' in dictionary
* Step 2: Checking inflection suffixes for 'dikenal'
* Step 3: Checking derivation suffixes for 'dikenal'
* Step 4b: Removing prefixes from 'dikenal'
* Removed prefix: 'dikenal' → 'kenal'
* Result: Found 'kenal' in dictionary

## Word: terdekat → dekat

* Step 1: Checking 'terdekat' in dictionary
* Step 2: Checking inflection suffixes for 'terdekat'
* Step 3: Checking derivation suffixes for 'terdekat'
* Step 4b: Removing prefixes from 'terdekat'
* Removed prefix: 'terdekat' → 'dekat'
* Result: Found 'dekat' in dictionary

## Word: kumpulan → kumpul

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

## Word: peter → ter

* Step 1: Checking 'peter' in dictionary
* Step 2: Checking inflection suffixes for 'peter'
* Step 3: Checking derivation suffixes for 'peter'
* Step 4b: Removing prefixes from 'peter'
* Removed prefix: 'peter' → 'ter'
* Result: Found 'ter' in dictionary

## Word: diadopsi → adopsi

* Step 1: Checking 'diadopsi' in dictionary
* Step 2: Checking inflection suffixes for 'diadopsi'
* Step 3: Checking derivation suffixes for 'diadopsi'
* Removed -i: 'diadopsi' → 'diadops'
* Restored original: suffix removal unsuccessful
* Step 4: Checking prefix-suffix combination: prefix='di', suffix='i'
* Step 4b: Removing prefixes from 'diadopsi'
* Removed prefix: 'diadopsi' → 'adopsi'
* Result: Found 'adopsi' in dictionary

## Word: tulisan → tulis

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

## Word: meluas → luas

* Step 1: Checking 'meluas' in dictionary
* Step 2: Checking inflection suffixes for 'meluas'
* Step 3: Checking derivation suffixes for 'meluas'
* Step 4b: Removing prefixes from 'meluas'
* Removed prefix: 'meluas' → 'luas'
* Result: Found 'luas' in dictionary

## Word: dianggap → anggap

* Step 1: Checking 'dianggap' in dictionary
* Step 2: Checking inflection suffixes for 'dianggap'
* Step 3: Checking derivation suffixes for 'dianggap'
* Step 4b: Removing prefixes from 'dianggap'
* Removed prefix: 'dianggap' → 'anggap'
* Result: Found 'anggap' in dictionary

## Word: xi → x

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

## Word: dialami → alami

* Step 1: Checking 'dialami' in dictionary
* Step 2: Checking inflection suffixes for 'dialami'
* Step 3: Checking derivation suffixes for 'dialami'
* Removed -i: 'dialami' → 'dialam'
* Restored original: suffix removal unsuccessful
* Step 4: Checking prefix-suffix combination: prefix='di', suffix='i'
* Step 4b: Removing prefixes from 'dialami'
* Removed prefix: 'dialami' → 'alami'
* Result: Found 'alami' in dictionary

## Word: terkena → kena

* Step 1: Checking 'terkena' in dictionary
* Step 2: Checking inflection suffixes for 'terkena'
* Step 3: Checking derivation suffixes for 'terkena'
* Step 4b: Removing prefixes from 'terkena'
* Removed prefix: 'terkena' → 'kena'
* Result: Found 'kena' in dictionary

## Word: acuan → acu

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

## Word: terdekat → dekat

* Step 1: Checking 'terdekat' in dictionary
* Step 2: Checking inflection suffixes for 'terdekat'
* Step 3: Checking derivation suffixes for 'terdekat'
* Step 4b: Removing prefixes from 'terdekat'
* Removed prefix: 'terdekat' → 'dekat'
* Result: Found 'dekat' in dictionary

## Word: terdekat → dekat

* Step 1: Checking 'terdekat' in dictionary
* Step 2: Checking inflection suffixes for 'terdekat'
* Step 3: Checking derivation suffixes for 'terdekat'
* Step 4b: Removing prefixes from 'terdekat'
* Removed prefix: 'terdekat' → 'dekat'
* Result: Found 'dekat' in dictionary

## Word: terdekat → dekat

* Step 1: Checking 'terdekat' in dictionary
* Step 2: Checking inflection suffixes for 'terdekat'
* Step 3: Checking derivation suffixes for 'terdekat'
* Step 4b: Removing prefixes from 'terdekat'
* Removed prefix: 'terdekat' → 'dekat'
* Result: Found 'dekat' in dictionary

## Word: ketiga → tiga

* Step 1: Checking 'ketiga' in dictionary
* Step 2: Checking inflection suffixes for 'ketiga'
* Step 3: Checking derivation suffixes for 'ketiga'
* Step 4b: Removing prefixes from 'ketiga'
* Removed prefix: 'ketiga' → 'tiga'
* Result: Found 'tiga' in dictionary

## Word: ketiga → tiga

* Step 1: Checking 'ketiga' in dictionary
* Step 2: Checking inflection suffixes for 'ketiga'
* Step 3: Checking derivation suffixes for 'ketiga'
* Step 4b: Removing prefixes from 'ketiga'
* Removed prefix: 'ketiga' → 'tiga'
* Result: Found 'tiga' in dictionary

## Word: ringkasan → ringkas

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

## Word: ketiga → tiga

* Step 1: Checking 'ketiga' in dictionary
* Step 2: Checking inflection suffixes for 'ketiga'
* Step 3: Checking derivation suffixes for 'ketiga'
* Step 4b: Removing prefixes from 'ketiga'
* Removed prefix: 'ketiga' → 'tiga'
* Result: Found 'tiga' in dictionary

# Final Stemmed Text:

penerapan metode k -nearest neighbors (k -nn) klasifikasi risiko diabetes dataset kesehatan susun oleh: emelsha viadra 152022056 naufal zaidan 152022168 prodi informatika fakultas teknologi industri institut teknologi nasional bandung 2024 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 dekat dipertimbangkan mengklasifikasikan data baru. pemilihan nilai k penting, k model peka data ( overfitting ), k model kehilangan sensitivitas pola lokal data. proses utama k -nn melibatkan perhitungan jarak data data dataset ada. jarak hitung metrik euclidean distance , manhattan distance , minkowski distance . euclidean distance salah metode digunakan, jarak ukur 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 kecil data baru. tahap akhir, algoritma menentukan kelas data berdasarkan mayoritas kelas k tetangga dekat tersebut. misa lnya, mayoritas k tetangga kelas diabetes positif, data diklasifikasikan diabetes positif. kelebihan utama k -nn kesederhanaannya, proses pelatihan rumit. namun, l emahannya algoritma lambat ukur dataset besar, menghitung jarak data data dataset. sejarah algoritma k-nearest neighbors (k-nn) akar perkembangan kecerdasan buat analisis data statistik. konsep dasar algoritma kenal abad -20, perkembangan signifikan teori klasifikasi pengenalan pola pertengahan 1900 -an. poin sejarah algoritma k -nn: 1. 1951: penemuan evelyn fix joseph hodges • algoritma k -nn kali diperkenalkan evelyn fix joseph hodges 1951 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 dekat kumpul data label. 2. 1967: penerapan pengenalan pola • 1967, algoritma k -nn populer buku berjudul "the theory of pattern recognition" thomas m. cover ter 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. 3. 1970 -an: perkembangan bidang pengklasifikasian • algoritma k -nn aplikasi pengklasifikasian, bidang statistik pembelajaran mesin. • k-nn adopsi pendekatan sederhana kuat klasifikasi data. aplikasi utamanya bidang bioinformatika, pengenalan wajah, pengenalan tulis tangan. 4. 1980 -1990 -an: penggunaan komputasi • era ini, k -nn populer munculnya komputasi cepat terjangkau. • penggunaan k -nn luas bidang kompleks komputer vision , pengenalan suara , diagnosis medis . 5. 2000 -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. 6. • k-nn anggap algoritma mudah dipahami diterapkan, relevan aplikasi pembelajaran mesin. • algoritma sederhana memiliki keterbatasan, rentan overfitting performa lambat dataset besar, 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 adalah: penjelasan: • x y data titik jaraknya dihitung. • x y nilai fitur -i titik x y. • n fitur dataset. contoh 1: diabetes salah penyakit alami masyarakat, deteksi mencegah komplikasi lanjut. studi ini, algoritma k -nearest neighbors (k -nn) mengklasifikasikan berisiko kena diabetes berdasarkan faktor kesehatan. tabel 1 .data sampel data sampel no gula darah insulin bmi usia diabet? 1 89 94 28,1 21 2 137 168 43,1 33 ya 3 78 88 31 26 ya 4 197 543 30,5 53 ya 5 189 846 30,1 59 ya 6 166 175 25,8 51 ya 7 118 230 45,8 31 ya 8 103 83 43,3 33 9 115 96 34,6 32 ya 10 126 235 39,3 27 11 143 146 36,6 51 ya 12 125 115 31,1 41 ya 13 97 140 23,2 22 14 145 110 22,2 57 15 158 245 31,6 28 ya 16 88 54 24,8 22 17 103 192 24 33 18 111 207 37,1 56 ya 19 180 70 34 26 20 171 240 45,4 54 ya 21 103 82 19,4 22 22 101 36 24,2 26 23 88 23 24,4 30 24 176 300 33,7 58 ya 25 150 342 34,7 42 26 100 71 38,5 26 27 110 125 32,4 27 penyelesaian : data k ita ambil 3 data acu pengujian data tabel 2. data uji data uji gula darah insulin bmi usia diabet? 187 304 37,7 41 ya 93 64 28,7 23 155 495 34 46 ya tabel 3. hasil pengujian data 1 data uji 1 jarak euclidian urut diabet? 232,80 17 145,22 10 242,50 22 239,62 21 542,36 27 131,62 8 101,99 6 236,63 19 220,31 16 93,17 5 164,32 11 199,02 14 188,59 12 199,74 15 67,29 3 ya 269,87 24 140,90 9 124,14 7 234,61 18 67,68 4 ya 238,82 20 282,18 25 298,43 26 21,02 1 ya 53,13 2 249,17 23 195,43 13 tabel 3. hasil pengujian data 2 data uji 2 jarak euclidian urut diabet? 30,34 7 114,28 16 28,55 5 ya 491,08 26 788,69 27 135,80 18 168,93 20 27,82 4 40,30 8 174,52 21 100,35 15 62,89 10 76,31 12 77,58 13 192,40 22 11,88 1 128,86 17 148,10 19 87,42 14 195,70 23 22,62 3 29,62 6 42,11 9 252,66 24 284,48 25 14,25 2 63,56 11 tabel 3. hasil pengujian data 3 data uji 3 jarak euclidian urut diabet? 407,21 19 327,88 11 414,71 20 64,26 1 ya 352,90 13 320,33 10 268,25 7 415,58 21 401,24 18 262,35 6 349,25 12 381,23 16 360,67 14 385,47 17 250,68 4 ya 446,80 25 307,87 9 291,53 8 426,20 23 255,88 5 ya 417,21 22 462,70 26 477,10 27 196,49 3 ya 153,14 2 428,04 24 373,21 15 pengujian ini, algoritma k -nearest neighbors (k -nn) k=5 menentukan risiko diabetes data uji berdasarkan jarak euclidean data sampel. data uji 1: tetangga dekat data uji hasil klasifikasi adalah: • jarak 21,02 data sampel no. 1 - "ya" • jarak 53,13 data sampel no. 2 - "tidak" • jarak 67,29 data sampel no. 3 - "ya" • jarak 67,68 data sampel no. 4 - "ya" • jarak 93,17 data sampel no. 5 - "tidak" tetangga terdekat, memiliki hasil "ya" memiliki hasil "tidak" . mayoritas hasil "ya", data uji diklasifikasikan berisiko diabetes . data uji 2: tetangga dekat data uji hasil klasifikasi adalah: • jarak 11,88 data sampel no. 1 - "tidak" • jarak 14,25 data sampel no. 2 - "tidak" • jarak 22,62 data sampel no. 3 - "tidak" • jarak 27,82 data sampel no. 4 - "tidak" • jarak 28,55 data sampel no. 5 - "ya" tetangga terdekat, memiliki hasil "tidak" memiliki hasil "ya". demikian, data uji diklasifikasikan berisiko diabetes . data uji 3: tetangga dekat data uji tiga hasil klasifikasi adalah: • jarak 64,26 data sampel no. 1 - "ya" • jarak 153,14 data sampel no. 2 - "tidak" • jarak 196,49 data sampel no. 3 - "ya" • jarak 250,68 data sampel no. 4 - "ya" • jarak 255,88 data sampel no. 5 - "ya" tetangga terdekat, memiliki hasil "ya" memiliki hasil "tidak" . itu, data uji tiga diklasifikasikan berisiko diabetes . ringkas k=5, hasil klasifikasi bahwa: • data uji tiga diklasifikasikan berisiko diabetes . • data uji diklasifikasikan berisiko diabetes .