

Sistem Deteksi Dini Diagnosa ISPA (Infeksi Saluran Pernapasan Akut) pada Anak dengan Metode Cosine Similarity

Denis Eka Cahyani
Jurusan Informatika
Universitas Sebelas Maret
Jalan Ir.Sutami 36 A
Surakarta 57126

Rini Anggrainingsih
Jurusan Informatika
Universitas Sebelas Maret
Jalan Ir.Sutami 36 A
Surakarta 57126

ABSTRAK

Infeksi Saluran Pernapasan Akut (ISPA) merupakan salah satu masalah kesehatan masyarakat yang utama pada bayi (0-11 bulan) dan balita (1-4 tahun) di Indonesia. Saat ini dibutuhkan pengenalan dini yang tepat untuk dapat mendeteksi gejala-gejala penyakit ISPA terutama yang terjadi pada anak supaya dapat mencegah keterlambatan dalam menangani penyakit ISPA. Salah satu peluang pengembangan teknologi deteksi dini untuk diagnosa penyakit ISPA ini yaitu dengan mengukur kesamaan (similaritas) gejala awal yang terjadi pada anak dengan gejala pada penyakit ISPA. Sistem deteksi dini diagnosa penyakit ISPA pada anak dengan metode cosine similarity bertujuan untuk mengukur efektifitas penggunaan metode cosine similarity dalam mendeteksi dini diagnosa penyakit ISPA pada anak. Pada penelitian ini metode cosine similarity dibandingkan dengan metode jaccard similarity. Penerapan metode pada sistem dilakukan dengan membuat representasi objek penelitian yaitu model data matriks dengan input jenis atribut data biner. Representasi objek ini digambarkan dengan tabel hubungan gejala yang diderita dengan jenis kemungkinan diagnosa penyakit pasien yang akan menjadi kamus data. Lalu kemudian diurutkan data gejala pada kamus data berdasarkan pada probabilitas tertinggi. Berdasarkan hasil pengujian yang telah dilakukan, diperoleh kesimpulan bahwa keefektifan metode cosine similarity bila dibandingkan dengan jaccard similarity adalah sama.

Kata Kunci

deteksi dini, similaritas gejala awal, diagnosa ISPA, cosine similarity, jaccard similarity

1. PENDAHULUAN

Infeksi Saluran Pernapasan Akut (ISPA) di Indonesia masih merupakan salah satu masalah kesehatan masyarakat yang utama pada bayi (0-11 bulan) dan balita (1-4 tahun). Diperkirakan kejadian ISPA pada balita di Indonesia yaitu sebesar 10-20% (Depkes RI, 2004). ISPA khususnya pneumonia banyak menyebabkan kematian pada balita. Berdasarkan Bryce, et al. (2005), Proportional Mortality Rate (PMR) yaitu ukuran yang berhubungan dengan jumlah kematian akibat kondisi tertentu untuk semua kematian dalam kelompok populasi yang sama untuk periode yang sama, pada balita karena pneumonia di dunia adalah sebesar 26%. Kemudian berdasarkan WHO (2005) dalam Depkes RI (2005) dikatakan bahwa PMR karena pneumonia untuk regional Asia Tenggara 2000-2003 adalah sebesar 19%.

ISPA di Indonesia menempati urutan pertama penyebab kematian pada kelompok bayi dan balita. ISPA juga berada pada daftar 10 penyakit terbanyak. Survey yang dilakukan menunjukkan bahwa ISPA merupakan penyebab kematian bayi terbesar di Indonesia dengan persentasi 22,30% dari seluruh kematian balita (Depkes RI, 2008). Kematian balita karena

ISPA secara nasional diperkirakan 6 orang per 1000 balita per tahun atau sekitar 150.000 balita per tahun (Depkes RI, 2002).

ISPA merupakan penyebab utama morbiditas (meratanya penyakit) dan mortalitas (ukuran jumlah kematian pada suatu populasi) penyakit menular di dunia. Hampir empat juta orang meninggal akibat ISPA setiap tahun, 98% -nya disebabkan oleh infeksi saluran pernafasan bawah. Tingkat mortalitas akibat ISPA pada bayi, anak, dan orang lanjut usia tergolong tinggi terutama di negara-negara dengan pendapatan per kapita rendah dan menengah. ISPA juga merupakan salah satu penyebab utama konsultasi atau rawat inap di sarana pelayanan kesehatan terutama pada bagian perawatan anak (WHO, 2007).

Saat ini dibutuhkan pengenalan dini yang tepat untuk dapat mendeteksi gejala-gejala penyakit ISPA terutama yang terjadi pada anak supaya dapat mencegah keterlambatan dalam menangani penyakit ISPA, karena kalau tidak dilakukan pengenalan dini untuk dapat mendeteksi penyakit ISPA bisa mengakibatkan dampak yang buruk bagi anak bahkan kematian. Gangguan kesehatan penyakit yang ringan merupakan salah satu petunjuk dari gangguan kesehatan yang lebih serius.

Teknologi deteksi dini untuk diagnosa penyakit ISPA masih memiliki peluang besar untuk dikembangkan. Salah satu peluang pengembangan teknologi deteksi dini untuk diagnosa penyakit ISPA ini yaitu dengan mengukur kesamaan (similaritas) gejala awal yang terjadi pada anak dengan gejala pada penyakit ISPA berdasar pada lokasi anatomik tubuh. Kesamaan antara gejala yang terjadi dengan gejala penyakit yang sebenarnya dapat ditunjukkan dengan fungsi similaritas (mengukur kesamaan) atau fungsi jarak (mengukur ketidaksamaan).

Beberapa fungsi similaritas dan fungsi jarak yang dapat dijumpai antara lain adalah Dice, Jaccard, Overlap, Asymmetric, Minowski Distance, Euclidean Distance, Correlation, dan Cosine. Untuk tujuan pengelompokan data jarak fungsi yang paling baik adalah dengan cosine similarity (Strehl, et al., 2000). Cosine similarity adalah metode similarity yang paling banyak digunakan untuk menghitung similaritas (kesamaan) diantara satu data dengan data lainnya (Tan, et al., 2006).

2. LANDASAN TEORI

2.1 Pengukuran Similarity

Saat ini dikembangkan banyak metode pengukuran similarity yang dapat digunakan untuk berbagai keperluan (Karhendana, 2008). Secara umum, fungsi similarity adalah fungsi yang menerima dua buah objek dan mengembalikan nilai kemiripan (similarity) antara kedua objek tersebut berupa bilangan riil. Umumnya, nilai yang dihasilkan oleh fungsi similarity berkisar pada interval [0...1]. Semakin besar hasil fungsi similarity, maka kedua objek yang dievaluasi dianggap semakin mirip.

Sebaliknya, semakin kecil hasil fungsi similarity, maka kedua objek tersebut dianggap semakin berbeda. Pada fungsi yang menghasilkan nilai pada jangkauan [0...1], nilai 1 melambangkan kedua objek persis sama, sedangkan nilai 0 melambangkan kedua objek sama sekali berbeda.

Umumnya, sebuah fungsi similarity antara dua objek (oi, oj) akan memenuhi sifat-sifat sebagai berikut (Karhendana, 2008) :

$$oi \ S : (oi, oi) > 0$$

$$oi,oj \ S : (oi, oj) > 0 \text{ (non negatif)}$$

$$oi,oj \ S : (oi, oj) = (oj, oi) \text{ (simetris)}$$

$$oi,oj \ S : (oi, oj) \ [0...1] \text{ (ternormalisasi)}$$

$$oi,oj,ok \ S : (oi, oj) + (oj, ok) > (oi, ok) \text{ (ketidaksamaan segitiga)}$$

Pengukuran similarity yang berbasis himpunan disebut juga koefisien asosiasi karena fungsi ini mengukur persamaan dan perbedaan antara dua objek pada himpunan atribut masing-masing objek. Beberapa fungsi similarity yang berbasis himpunan adalah sebagai berikut (Karhendana, 2008) :

Koefisien Jaccard

$$(oi,oj) = (|oi \cap oj|) / (|oi \cup oj|) \quad (1)$$

Koefisien Jaccard menghitung similarity antara dua objek, X dan Y yang dinyatakan dalam dua vektor (Tan et al, 2006). Contohnya sebagai berikut:

$$X = (1,1,0,0,1,0,1,1,0) \ Y = (1,1,1,0,0,0,1,1,1)$$

$$X.Y = (1.1) + (1.1) + (0.1) + (0.0) + (1.0) + (0.0) + (1.1) + (1.1) + (0.1) = 4$$

$$X^2 = 12+12+02 +02+12+02+12+12+02 = 5$$

$$Y^2 = 12+12+12+02+02+02+12+12+12 = 6$$

$$J(x, y) = (X.Y) / (X.X + Y.Y - X.Y)$$

$$J(x, y) = 4 / (5 + 6 - 4) = 4/7 = 0,571$$

Koefisien Cosine

$$(oi,oj) = (|oi \cap oj|) / (|oi| \cdot |oj|) \quad (2)$$

Koefisien Cosine menghitung similarity antara dua objek, X dan Y yang dinyatakan dalam dua vektor, contohnya sebagai berikut:

$$X = (1,1,0,0,1,0,1,1,0) \ Y = (1,1,1,0,0,0,1,1,1)$$

$$x.y = (1.1) + (1.1) + (0.1) + (0.0) + (1.0) + (0.0) + (1.1) + (1.1) + (0.1) = 4$$

$$|x| = 12+12+02 +02+12+02+12+12+02$$

$$= 5 = 2,236$$

$$|y| = 12+12+12+02+02+02+12+12+12$$

$$= 6 = 2,449$$

$$C(X,Y) = (X.Y) / (|X| \cdot |Y|)$$

$$C(x,y) = 4 / (2,236 \cdot 2,449) = 4/5,475 = 0.730$$

Fungsi-fungsi similarity tersebut hanya berlaku untuk data matriks yang atributnya berjenis biner (0 atau 1). Untuk diterapkan pada atribut data kontinyu yang bernilai riil, fungsi-fungsi tersebut harus digeneralisasi. Sebab fungsi-fungsi similarity tersebut tidak bisa digunakan pada perhitungan similarity data geometrik. Namun apabila sudah digeneralisasi,

fungsi tersebut dapat digunakan untuk perhitungan similarity data geometrik.

Koefisien cosine merupakan pengukuran similarity yang paling sering digunakan untuk keperluan pengelompokan data. Cosine similarity ini sebanding dengan sudut antara dua vektor data dan tidak terpengaruh oleh panjang data (Karhendana, 2008).

Sedangkan jaccard similarity juga merupakan metode yang dipakai untuk menghitung similarity antara dua objek. Koefisien jaccard menghitung similarity antara dua objek, X dan Y yang dinyatakan dalam dua buah vektor (Tan, et al., 2006).

Cosine Similarity merupakan ukuran kesamaan antara dua elemen dengan mengukur cosinus dari sudut antara mereka. Cosinus 0 adalah 1 dan kurang dari 1 untuk setiap sudut lainnya dan nilai terendah dari cosinus adalah -1. Cosinus dari sudut antara dua vektor menentukan apakah dua vektor tersebut menunjuk ke arah yang sama. Cosine memiliki sifat unik dimana nilainya diperbesar pada sudut 00 sampai 450, sehingga selain memiliki sifat yang bisa digunakan untuk data yang bersifat dekat dan hampir sama, cosine juga cocok untuk vektor data yang bersifat jarang (sparse) dan similarity-nya cenderung rendah (Karhendana, 2008).

2.2 Jenis Atribut Data

Pengukuran analisis cluster yang digunakan mempunyai jenis atribut. Jenis atribut data yang umum digunakan adalah (Karhendana, 2008):

1. Biner : jenis atribut ini hanya memiliki dua kemungkinan nilai, yaitu 0 dan 1 atau benar dan salah.
2. Diskrit : jenis atribut ini berupa nilai yang diskrit, misalnya bilangan bulat. Contohnya adalah hasil perhitungan banyaknya item.
3. Kontinyu : jenis atribut ini berupa nilai yang kontinyu dalam bilangan riil. Contohnya adalah hasil pengukuran berat benda.

2.3 Probabilitas

Probabilitas mempunyai banyak persamaan seperti kemungkinan, kesempatan dan kecenderungan. Probabilitas menunjukkan kemungkinan terjadinya suatu peristiwa yang bersifat acak. Suatu peristiwa disebut acak jika terjadinya peristiwa tersebut tidak diketahui sebelumnya. Oleh karena itu, probabilitas dapat digunakan sebagai alat ukur terjadinya peristiwa di masa yang akan datang.

Nilai probabilitas yang paling kecil adalah 0 yang berarti bahwa peristiwa tersebut pasti tidak akan terjadi. Sedangkan nilai probabilitas yang terbesar adalah 1 yang berarti bahwa peristiwa tersebut pasti akan terjadi. Secara umum, nilai probabilitas suatu peristiwa A adalah :

$$0 < P(A) < 1$$

Definisi mengenai probabilitas dapat dilihat dari tiga macam pendekatan. Salah satunya yaitu dengan pendekatan klasik. Menurut pendekatan klasik, probabilitas didefinisikan sebagai hasil bagi banyaknya peristiwa yang dimaksud dengan seluruh peristiwa yang mungkin. Dirumuskan sebagai berikut (Budiarto, 2002):

$$P(A) = \frac{n(A)}{n(S)} \quad (3)$$

Dimana P(A) = probabilitas terjadinya peristiwa A

n(A) = jumlah peristiwa A

n(S) = jumlah keseluruhan peristiwa yang mungkin terjadi.

2.4 Infeksi Saluran Pernapasan Akut (ISPA)

ISPA merupakan singkatan dari Infeksi Saluran Pernafasan Akut, istilah ini diadaptasi dari istilah dalam Bahasa Inggris Acute Respiratory Infections (ARI). ISPA ini merupakan penyakit infeksi akut yang menyerang salah satu bagian atau lebih dari saluran napas mulai dari hidung (saluran pernapasan atas) sampai alveoli (saluran pernapasan bawah) termasuk jaringan adneksanya seperti sinus rongga telinga tengah dan pleura (Depkes RI, 2001).

Istilah ISPA meliputi tiga unsur yaitu: Infeksi, Saluran Pernafasan dan Akut, dengan pengertian sebagai berikut (Depkes RI, 2004) :

- Infeksi

Infeksi adalah masuknya kuman atau mikroorganisme ke dalam tubuh manusia dan berkembang biak sehingga menimbulkan gejala penyakit.

- Saluran Pernafasan

Saluran pernafasan adalah organ mulai dari hidung hingga alveoli beserta organ adneksanya seperti sinus-sinus, rongga telinga tengah dan pleura. ISPA secara anatomis mencakup saluran pernafasan bagian atas, saluran pernafasan bagian bawah (termasuk jaringan paru-paru) dan organ adneksa saluran pernafasan. Dengan batasan ini, jaringan paru termasuk dalam saluran pernafasan (respiratory tract).

- Akut

Infeksi akut adalah infeksi yang berlangsung sampai dengan 14 hari. Batas 14 hari diambil untuk menunjukkan proses akut meskipun untuk beberapa penyakit yang dapat digolongkan dalam ISPA proses ini dapat berlangsung lebih dari 14 hari.

Penyebab terjadinya ISPA adalah virus, bakteri dan jamur. Namun kebanyakan disebabkan oleh virus (Purnomo, 2008). Pemberian air susu ibu (ASI) eksklusif dan lingkungan menjadi faktor yang mempengaruhi kejadian ISPA (Abbas, 2011)

Infeksi Saluran Pernafasan Akut (ISPA) menjadi dua bagian, yaitu: Infeksi Saluran Pernafasan atas Akut (ISPAa) dan Infeksi Saluran Pernafasan bawah Akut (ISPAb). (Widjaja, 2003)

Infeksi Saluran Pernafasan Atas dalam Bahasa Indonesia dikenal sebagai ISPAa (Infeksi Saluran Pernafasan atas Akut) atau ISPA (Infeksi Saluran naPas Atas) atau URI dalam Bahasa Inggris adalah penyakit infeksi akut yang melibatkan organ saluran pernafasan, hidung, sinus, faring, atau laring. Infeksi Saluran Pernafasan atas Akut terdiri dari: Rhinitis, Sinusitis, Otitis Media (Radang Telinga Tengah), Stomatitis (mulut), Gingivitis (gusi), Faringitis (Radang Tenggorokan/ amandel). Sedangkan jenis Infeksi Saluran Pernafasan bawah Akut terdiri dari Infeksi Epiglottis (pita suara), Bronchitis, dan Pneumonia.

ISPA juga mempunyai kemiripan gejala dengan beberapa penyakit lainnya antara lain dengan penyakit asma dan TBC. Penyakit asma berasal dari kata "Asthma" yang diambil dari bahasa Yunani yang berarti "sukar bernapas". Asma adalah penyakit kronis (berlangsung lama) yang ditandai oleh sesak napas disertai bunyi ngik-ngik (mengi) atau batuk persisten dimana derajat keparahan setiap orang berbeda-beda (Prasetyo, 2010). Penyakit asma dikenal karena adanya gejala sesak napas, batuk dan mengi yang disebabkan oleh penyempitan saluran napas. Gejala awal penyakit asma ini mirip dengan gejala penyakit ISPA.

Penyakit TBC juga merupakan penyakit yang mempunyai gejala awal mirip dengan ISPA. Tuberkulosis adalah penyakit menular langsung yang disebabkan oleh kuman TB (*Mycobacterium tuberculosis*). Sebagian besar kuman TB menyerang paru, tetapi dapat juga mengenai organ tubuh lainnya. TBC lebih sering menyerang paru-paru, namun juga dapat menyerang bagian tubuh lain seperti selaput otak, kulit, tulang, kelenjar getah bening, dan bagian tubuh lainnya (Smeltzer, 2002).

2.5 Skala Likert

Skala Likert menurut Djaali (2008) ialah skala yang dapat dipergunakan untuk mengukur sikap, pendapat, dan persepsi seseorang atau sekelompok orang tentang suatu gejala. Skala Likert adalah suatu skala psikometrik yang umum digunakan dalam kuesioner, dan merupakan skala yang paling banyak digunakan dalam riset berupa survei. Biasanya disediakan lima pilihan skala dengan format seperti:

1. Sangat tidak setuju
2. Tidak setuju
3. Netral
4. Setuju
5. Sangat setuju

Selain pilihan dengan lima skala seperti contoh di atas, kadang digunakan juga skala dengan tiga, tujuh atau sembilan tingkat. Skala likert digunakan untuk menggambarkan secara kasar posisi individu dalam kelompoknya (posisi relatif), ingin membandingkan skor subyek dengan kelompok normatifnya dan ingin menyusun skala pengukuran yang sederhana dan mudah dibuat.

3. METODOLOGI PENELITIAN

Metodologi penelitian diuraikan ke dalam skema tahapan penelitian untuk memberikan petunjuk yang jelas, teratur, dan sistematis seperti yang ditunjukkan seperti gambar 1.

3.1 Pengumpulan Data

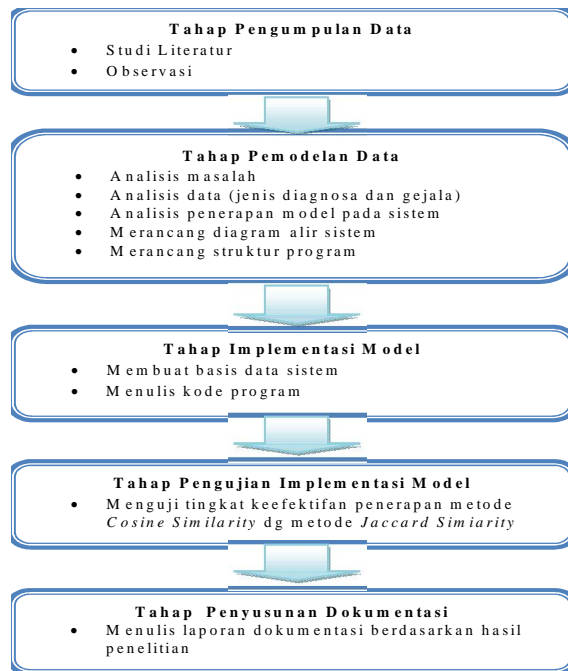
Tahap pertama yang dilakukan dalam penelitian ini adalah pengumpulan data. Tahap ini terbagi menjadi dua yaitu studi literatur dan observasi.

3.1.1 Studi Literatur

Studi literatur dilakukan melalui dua cara yaitu penelusuran internet dan membaca buku-buku untuk mendapatkan informasi tentang penelitian yang relevan dengan objek yang dikaji ini guna memperoleh ketepatan langkah dalam pelaksanaan penelitian. Selain itu juga untuk mengumpulkan bahan materi untuk melakukan penelitian seperti materi mengenai ISPA (Infeksi Saluran Pernapasan Akut) atau algoritma cosine similarity.

3.1.2 Observasi

Observasi dilakukan melalui pengamatan langsung di praktek dokter, mengumpulkan data sekunder berupa data pemeriksaan ISPA yang dibutuhkan serta melakukan wawancara dengan dokter tersebut. Pada penelitian ini dilakukan pengamatan langsung di tempat praktek dr. Safari Wiji Jatmiko dan wawancara dengan dr. Iwan Setiawan, Sp.THT.



Gambar 1. Alur Metodologi Penelitian

3.2 Pemodelan Data

3.2.1 Analisis Masalah

Permasalahan yang dianalisa adalah bagaimana mengukur seberapa besar tingkat keefektifan metode cosine similarity dalam mendeteksi dini diagnosa penyakit ISPA (Infeksi Saluran Pernapasan Akut) pada anak bila dibandingkan dengan metode jaccard similarity.

3.2.2 Analisis Data (jenis diagnosa dan gejala)

Analisa data diperlukan untuk mengetahui bagaimana langkah untuk mengolah data menggunakan metode yang dipakai dalam penelitian ini. Data yang digunakan untuk mendeteksi dini penyakit ISPA ini adalah data jenis-jenis diagnosa penyakit ISPA serta penyakit yang mempunyai gejala awal mirip dengan penyakit ISPA. Lalu data gejala-gejala yang dialami pasien ketika menderita masing-masing jenis dari diagnosa penyakit ISPA tersebut.

3.2.3 Analisis Penerapan Model pada Sistem

Dalam menerapkan metode cosine similarity pada sistem yang akan dibangun, langkah-langkah yang dilakukan yaitu pertama dengan membuat representasi objek penelitian dengan model data matriks dengan input jenis atribut data yaitu biner. Jenis atribut data ini hanya memiliki dua kemungkinan nilai yaitu 0 dan 1 atau tidak terjadi gejala dan muncul gejala pada penderita. Representasi objek ini digambarkan dengan tabel hubungan gejala yang diderita dengan jenis kemungkinan diagnosa penyakit pasien yang nantinya akan menjadi kamus data dalam sistem yang akan dibangun.

Contoh :

Tabel 1. Tabel contoh analisa penerapan model pada sistem

Gejala \ Penyakit	Infeksi 001
G001	1
G002	0

1 = muncul gejala G001 dimana gejala tersebut merupakan salah satu tanda dari Infeksi 001

0 = tidak muncul gejala G002 dimana gejala tersebut tidak merupakan salah satu tanda dari infeksi 001

Lalu yang kedua dengan mengurutkan data gejala pada kamus data berdasarkan probabilitas tertinggi, yaitu kemungkinan gejala yang paling sering terjadi pada infeksi yang diderita oleh pasien. Sehingga diperoleh data dari urutan paling atas gejala yang paling sering terjadi atau muncul pada infeksi. Urutan data gejala dapat digambarkan dengan tabel urutan kemungkinan gejala yang sering terjadi pada infeksi yang diderita pasien.

3.2.4 Merancang diagram alir sistem

Membuat rancangan diagram alir yang digunakan dalam sistem yang ditunjukkan pada Gambar 2.

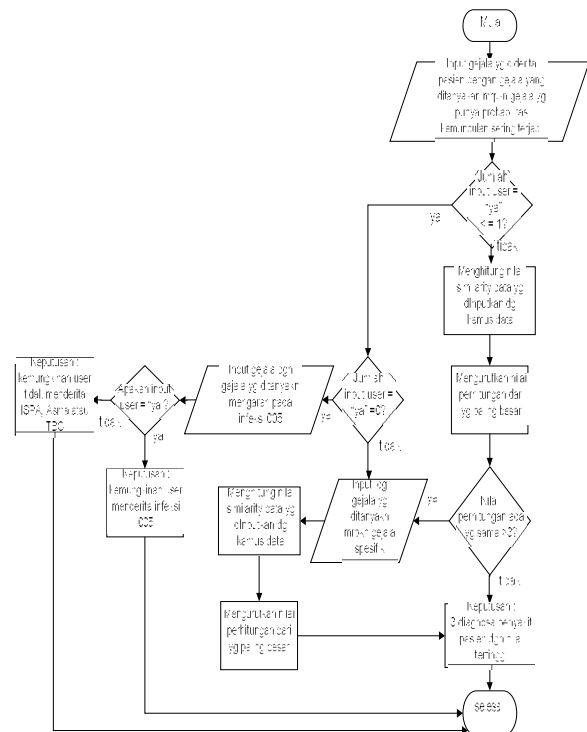
3.2.5 Merancang struktur program

Merancang desain dan struktur basis data serta struktur program sistem deteksi dini diagnosa ISPA pada anak dengan menggunakan metode cosine similarity. Metode cosine similarity dapat ditunjukkan sama dengan persamaan 2.

Pada penelitian ini juga merancang struktur program sistem dengan metode jaccard similarity. Metode jaccard similarity dapat ditunjukkan sama dengan persamaan 1.

3.3 Implementasi Model

Implementasi sistem deteksi dini diagnosa penyakit ISPA pada anak dengan metode cosine similarity menggunakan database MYSQL dan bahasa pemrograman PHP dan java mobile android.



Gambar 2. Diagram alir rancangan sistem

3.4 Pengujian Implementasi Model

Pada tahap ini dilakukan pengujian terhadap penerapan metode pada sistem. Data gejala yang diderita pasien diujikan dalam sistem deteksi dini diagnosa ISPA sehingga diperoleh diagnosa penyakit dari sistem tersebut. Sistem akan memberikan rekomendasi diagnosa penyakit sebanyak tiga diagnosa dengan nilai perhitungan yang paling tinggi. Hal ini mengacu pada pengukuran berdasarkan skala likert dengan menggunakan tiga tingkat. Rekomendasi diagnosa penyakit dipilih dengan menggunakan tiga tingkat karena merupakan tingkat yang paling kecil dalam skala likert sehingga jumlah rekomendasi diagnosa penyakit tidak terlalu banyak. Skala likert digunakan untuk sebagai gambaran posisi diagnosa jenis penyakit ISPA terhadap semua jenis penyakit ISPA dalam penelitian ini.

Dalam penelitian ini diujikan data pasien sebanyak 29 buah. Tahap pengujian dengan sistem dilakukan menggunakan dua metode yang berbeda yaitu metode cosine similarity dan jaccard similarity sebagai metode pembandingan.

Untuk dapat mengukur besar tingkat keefektifan metode cosine similarity jika dibandingkan jaccard similarity diperlukan beberapa parameter untuk diujikan. Keefektifan dapat diartikan tepat guna atau tepat sasaran, yang berarti dalam penelitian ini diujikan metode mana yang lebih tepat dalam mencapai tujuan. Parameter untuk mengukur keefektifan dalam penelitian ini dapat dilihat dari keakuratan hasil diagnosa sistem tersebut.

Keakuratan hasil diagnosa bisa dilihat dengan membandingkan hasil diagnosa dari sistem dengan diagnosa yang diperoleh dokter, sehingga akan diperoleh hasil diagnosa sistem dengan menggunakan metode mana yang lebih akurat untuk dapat mendeteksi dini diagnosa penyakit ISPA pada anak. Hasil pengujian berdasarkan keakuratan data dapat diperoleh dari rumus di bawah ini.

$$\text{Akurasi} = 100\% \times \frac{\text{jumlah data yang akurat}}{\text{total jumlah data}} \quad (4)$$

Kemudian keakuratan juga bisa dilihat dari membandingkan urutan hasil diagnosa sistem antara metode cosine similarity dengan jaccard similarity. Besarnya nilai perbedaan urutan hasil dapat diperoleh dari rumus di bawah ini

$$\text{Urutan berbeda} = 100\% \times \frac{\text{jumlah data yang berbeda urutan}}{\text{total jumlah data}}$$

Hasil pengujian data dapat dianalisis berdasarkan nilai presisi. Pengujian dengan nilai presisi ini bertujuan untuk mengetahui besarnya kemungkinan kesalahan dalam mendiagnosa pada sistem. Nilai presisi ini diperoleh dari rata-rata nilai *similarity* data yang akurat dalam penelitian ini.

$$\text{Nilai presisi} = 100\% \times \frac{\sum (n \times \text{sim})}{\text{total jumlah data}} \quad (6)$$

Keterangan :

n= jumlah data yang akurat pada urutan ke -i

sim = rata-rata nilai similarity data yang akurat urutan ke -i

4. PENGUJIAN

Tahap pengujian bertujuan untuk mengetahui keefektifan metode cosine similarity dalam mendeteksi dini diagnosa ISPA pada anak. Dalam penelitian ini juga akan mengukur keefektifan metode jaccard similarity dalam mendeteksi dini diagnose ISPA pada anak sebagai pembandingan metode similarity. Metode jaccard similarity merupakan salah satu metode yang digunakan untuk menghitung similari tas (kesamaan) antara data satu dengan data lainnya.

Pengujian ini dilakukan dengan menguji keakuratan hasil diagnosa sistem yang dibandingkan hasil diagnosa dokter dengan input gejala yang diderita oleh pasien. Sehingga dari hasil pengujian tersebut akan diperoleh keakuratan hasil diagnosa yang benar di urutan rekomendasi ke berapa dalam sistem tersebut. Setelah diujikan semua data dapat dihitung besarnya keakuratan masing-masing pengujian metode sehingga dapat dibandingkan dengan metode mana yang lebih akurat untuk dapat mendeteksi dini diagnosa penyakit ISPA pada anak.

Pengujian pertama dilakukan dengan mengujikan data gejala pasien dalam sistem deteksi dini menggunakan metode cosine similarity. Input dari pengujian ini adalah data gejala pasien sebanyak 29 data. Hasil dari pengujian ini diharapkan memperoleh diagnosa sistem dari urutan pertama hingga urutan ketiga. Hasil diagnosa sistem tersebut lalu dibandingkan dengan diagnosa dokter sehingga diperoleh keakuratan hasil diagnosa sistem tersebut.

Pengujian kedua dilakukan dengan mengujikan data gejala pasien dalam sistem deteksi dini menggunakan metode jaccard similarity. Input dari pengujian ini adalah data gejala pasien sebanyak 29 data. Hasil dari pengujian ini diharapkan memperoleh diagnosa sistem dari urutan pertama hingga urutan ketiga. Hasil diagnosa sistem tersebut lalu dibandingkan dengan diagnosa dokter sehingga diperoleh keakuratan hasil diagnosa sistem tersebut.

Hasil pengujian sistem dengan metode cosine similarity menunjukkan bahwa terdapat 19 data yang betul di urutan ke-1, 1 data betul di urutan ke-2, 5 data betul di urutan ke-3 dan 4 data yang salah atau tidak akurat dari 29 data yang diujikan. Sedangkan untuk hasil pengujian sistem dengan metode jaccard similarity menunjukkan bahwa juga terdapat 19 data yang betul di urutan ke-1, 1 data betul di urutan ke-2, 5 data betul di urutan ke-3 dan 4 data yang salah atau tidak akurat dari 29 data yang diujikan.

Keakuratan hasil diagnosa juga bisa dilihat dari membandingkan urutan hasil diagnosa sistem antara metode cosine similarity dengan jaccard similarity. Perbandingan urutan hasil diagnosa dengan metode cosine similarity dan jaccard similarity dalam mendeteksi dini diagnosa ISPA pada anak dapat dilihat pada Tabel 2.

Tabel 2. Tabel perbandingan metode cosine similarity & jaccard similarity

	<i>Cosine similarity</i>			<i>Jaccard similarity</i>			Perbandingan
	Urutan ke-1	Urutan ke-2	Urutan ke-3	Urutan ke-1	Urutan ke-2	Urutan ke-3	
Data 1	100% Otitis media	63% pneumonia	58% sinusitis	100% Otitis media	40% pneumonia	33% sinusitis	sama
Data 2	71% Rhinitis	41% Asma	35% TBC	50% Rhinitis	25% Asma	20% TBC	sama
Data 3	71% Otitis media	67% Pneumonia	67% faringitis	42% Pneumonia	40% Otitis media	36% faringitis	Beda, urutan ke-1
Data 4	80% Faringitis	73% sinusitis	54% bronchitis	67% Faringitis	57% sinusitis	38% bronchitis	sama
Data 5	86% asma	75% TBC	67% pneumonia	75% asma	60% TBC	50% pneumonia	sama
Data 6	81% Asma	70% TBC	63% pneumonia	67% Asma	50% TBC	40% pneumonia	sama
Data 7	57% tbc	57% Rhinitis	51% pneumonia	40% TBC	33% rhinitis	33% pneumonia	sama
Data 8	61% Sinusitis	57% asma	50% stomatitis	43% Sinusitis	40% asma	29% faringitis	Beda, urutan ke-3
Data 9	73% Sinusitis	63% otitis media	60% pneumonia	57% Sinusitis	43% otitis media	43% pneumonia	sama
Data 10	87% TBC	77% pneumonia	47% bronchitis	75% TBC	60% pneumonia	29% bronchitis	sama
Data 11	Tidak menderita ispa, asma atau tbc			Tidak menderita ispa, asma atau tbc			sama
Data 12	77% Sinusitis	67% Pneumonia	67% faringitis	63% Sinusitis	50% Pneumonia	50% faringitis	sama
Data 13	73% Pneumonia	66% sinusitis	61% TBC	57% Pneumonia	50% sinusitis	43% TBC	sama
Data 14	70% Rhinitis	40% asma	35% TBC	50% Rhinitis	25% asma	20% TBC	sama
Data 15	Tidak menderita ispa, asma atau tbc			Tidak menderita ispa, asma atau tbc			sama
Data 16	81% rhinitis	52% pneumonia	52% faringitis	67% rhinitis	33% pneumonia	33% faringitis	sama
Data 17	71% otitis media	67% pneumonia	61% sinusitis	50% otitis media,	50% pneumonia	43% sinusitis	sama
Data 18	86% asma	75% TBC	70% rhinitis	75% asma	60% TBC	50% rhinitis	sama
Data 19	91% sinusitis	63% otitis media	60% faringitis	83% sinusitis	43% faringitis	40% otitis media	Beda, urutan ke-2
Data 20	81% bronchitis	67% pneumonia	67% faringitis	67% bronchitis	50% pneumonia	50% faringitis	sama

Tabel 2 membandingkan antara urutan pertama sampai ke tiga diagnosa sistem menggunakan metode cosine similarity dan urutan pertama sampai ketiga diagnosa sistem menggunakan metode jaccard similarity. Hasil dari perbandingan tersebut adalah terdapat tiga data yang berbeda urutannya yaitu data 3 berbeda urutan pertama, data 8 berbeda urutan ketiga, dan data 19 berbeda urutan kedua dari 29 data yang dibandingkan.

5. ANALISIS HASIL

Berdasarkan hasil pengujian yang telah dilakukan, maka bisa dianalisis perbandingan keakuratan data antara metode cosine similarity dengan jaccard similarity. Tabel perbandingan keakuratan dapat ditunjukkan dengan Tabel 3.

Tabel 3. Tabel perbandingan keakuratan data

No	Jenis Metode	Jumlah data	Urutan ke-1	Urutan ke-2	Urutan ke-3
1	<i>Cosine similarity</i>	29	66% (19 data)	3,4% (1 data)	17% (5 data)

Perbandingan tingkat keakuratan sistem antara metode cosine similarity dengan jaccard similarity adalah sama. Tingkat keakuratan sistem dalam mendeteksi dini diagnosa penyakit ISPA pada anak sebesar 86%. Nilai ini diperoleh dari jumlah nilai keakuratan sistem yang berada di urutan ke-1 sampai urutan ke-3. Tingkat kesalahan keakuratan sistem dalam mendiagnosa sistem yaitu sebesar 14%. Penyakit TBC tidak bisa muncul dalam hasil keputusan diagnosa dokter karena untuk saat ini untuk dapat mendiagnosa penyakit TBC harus ditunjang dengan menggunakan alat kedokteran.

Perbandingan metode cosine similarity dan jaccard similarity dalam mendeteksi dini diagnosa ISPA pada anak berdasarkan urutan mendiagnosa penyakit telah digambarkan pada Tabel 2.

Dari data yang diperoleh perbandingan urutan diagnosa sistem antara metode cosine similarity dengan jaccard similarity berbeda sebesar 10%. Namun hasil akhir diagnosa dari diagnosa ke dua metode adalah sama. Oleh karena itu walaupun berbeda urutannya sebesar 10% namun hal ini tidak mempengaruhi hasil

keakuratan data karena hasil akhir dari diagnosa sistem kedua metode ini sama.

Hasil pengujian data dapat dianalisis berdasarkan nilai presisi. Pengujian dengan nilai presisi ini bertujuan untuk mengetahui besarnya kemungkinan kesalahan dalam mendiagnosa pada sistem.

Perbandingan nilai presisi data ditunjukkan dengan Tabel 4.

Tabel 4. Tabel perbandingan nilai presisi data

No	Jenis Metode	Jumlah data	Urutan ke-1	Urutan ke-2	Urutan ke-3	Urutan ke-4	Urutan ke-5	Urutan ke-6	Nilai presisi
1	Cosine similarity	29	73% (19 data)	77% (1 data)	61% (5 data)	64% (2 data)	52% (1 data)	33% (1 data)	68%
2	Jaccard similarity	29	54% (19 data)	60% (1 data)	39% (5 data)	47% (2 data)	33% (1 data)	20% (1 data)	50%

Hasil nilai urutan ke-1 dari tabel di atas diperoleh dari rata-rata nilai similarity urutan ke-1 dimana keakuratannya betul (sama dengan diagnosa dokter). Sedangkan hasil nilai urutan ke-2 juga diperoleh dari rata-rata nilai similarity urutan ke-2 dimana keakuratannya betul, begitu juga dengan nilai urutan ke-6.

Tabel 4 menunjukkan bahwa nilai presisi data dengan metode cosine similarity sebesar 0.68 (68%). Artinya kemungkinan kesalahan diagnosa pada sistem dengan metode cosine similarity sebesar 32%. Hal ini diperoleh dari anggapan nilai untuk diagnosa dokter sebesar 100% dikurangi nilai presisi sistem dengan metode cosine similarity.

Sedangkan nilai presisi data dengan metode jaccard similarity sebesar 0.50 (50%). Artinya kemungkinan kesalahan diagnosa pada sistem dengan metode jaccard similarity sebesar 50%. Hal ini diperoleh dari anggapan nilai untuk diagnosa dokter sebesar 100% dikurangi nilai presisi sistem dengan metode jaccard similarity.

Dari hasil analisa tersebut dapat dilihat nilai presisi data dengan metode cosine similarity lebih tinggi dibandingkan jaccard similarity. Namun hal ini tidak mempengaruhi besarnya keefektifan metode karena sifat dari cosine similarity yang besar nilainya selalu lebih tinggi daripada jaccard similarity, hasil nilai presisi ini dapat digunakan untuk mendukung hasil penelitian ini.

Dari berbagai pembahasan di atas, dapat disimpulkan bahwa tingkat keakuratan metode cosine similarity dalam mendeteksi dini diagnosa penyakit ISPA pada anak adalah sebesar 0.86 (86%) sedangkan metode jaccard similarity juga sebesar 0.86 (86%) sehingga dapat disimpulkan bahwa keefektifan metode cosine similarity bila dibandingkan dengan jaccard similarity adalah sama.

Hal ini juga didukung dengan perbandingan berdasarkan urutan diagnosa pada kedua metode ini mempunyai perbedaan sebesar 10% namun hal ini tidak mempengaruhi hasil keakuratan data karena hasil akhir dari diagnosa sistem kedua metode ini adalah sama.

6. KESIMPULAN

Kesimpulan akhir berdasarkan penelitian yang telah dilakukan adalah bahwa hasil pengujian metode yang telah dilakukan menunjukkan bahwa tingkat keefektifan metode cosine similarity dalam mendeteksi dini diagnosa penyakit ISPA pada anak adalah sebesar 0.86 (86%) sedangkan metode jaccard similarity sebesar juga 0.86 (86%) sehingga dapat disimpulkan bahwa keefektifan metode cosine similarity bila dibandingkan dengan jaccard similarity adalah sama.

Hal ini juga didukung dengan hasil pengujian berdasarkan nilai presisi. Nilai presisi data dengan metode cosine similarity

sebesar 0.68 (68%). Sedangkan nilai presisi data dengan metode jaccard similarity sebesar 0.50 (50%). Hal ini menunjukkan bahwa nilai presisi data dengan metode cosine similarity lebih tinggi dibandingkan jaccard similarity. Serta perbandingan berdasarkan urutan diagnosa pada kedua metode ini mempunyai perbedaan sebesar 10% namun hal ini tidak mempengaruhi hasil keakuratan data karena hasil akhir dari diagnosa sistem kedua metode ini adalah sama.

Saran untuk pengembangan sistem serupa dimasa mendatang:

1. Sebaiknya data observasi yang digunakan dilengkapi jenis-jenis penyakit ISPA dan observasi yang lebih cermat untuk menghasilkan hasil pengujian yang lebih jelas dan akurat.
2. Data penyakit mirip ISPA yang digunakan dalam sistem lebih banyak lagi sehingga sistem dapat lebih akurat dalam mendeteksi dini penyakit ISPA.

7. DAFTAR PUSTAKA

- [1] Abbas, Pujiati. Haryati, Aprillia Sri. 2011 . Hubungan Pemberian ASI Eksklusif dengan Kejadian Infeksi Saluran Pernapasan Akut (ISPA) pada Bayi. *Majalah Ilmiah Sultan Agung* Vol. 123, No.49
- [2] Bryce J, Victora CG, Habicht JP, Black RE and Scherpbier RW. 2005. Programmatic pathways to child survival: results of a multi-country evaluation of Integrated Management of Childhood Illness. *Health Policy Plan* 20 Supplement 1
- [3] Budiarto, Eko. 2002. *Biostatistika untuk Kedokteran dan Kesehatan Masyarakat*. Jakarta: Buku Kedokteran EGC.
- [4] Depkes RI. 2001. *Pedoman Pemberantasan Penyakit Infeksi Saluran Pernafasan Akut*. Jakarta: Dirjen PPM & PLP
- [5] Depkes RI. 2002. *Pedoman Pemberantasan Penyakit Infeksi Saluran Pernapasan Akut untuk Penanggulangan Pneumonia Balita*. Jakarta: Dirjen PPM & PLP
- [6] Depkes RI. 2004. *Angka Kematian Bayi Masih Tinggi. ISPA Pembunuh Utama*. Jakarta: Dirjen PPM & PLP
- [7] Depkes RI. 2004. *Pengertian ISPA dan Pneumonia*. Jakarta: Dirjen PPM & PLP
- [8] Depkes RI. 2005. *Profil Kesehatan Tahun 2005*. Jakarta: Dirjen PPM & PLP
- [9] Depkes RI. 2008. *Profil Kesehatan di Indonesia*. Jakarta: Dirjen PPM & PLP
- [10] Karhendana, Arie. 2008. *Pemanfaatan Document Clustering pada Agregator Berita*. Tugas Akhir. Program Studi Teknik Informatika, Institut Teknologi Bandung.
- [11] Prasetyo, Budi. 2010. *Seputar Masalah Asma*. Jakarta: Divapress.
- [12] Purnomo, Wahyu. 2008. Hubungan antara Pengetahuan dan Sikap Ibu dengan Upaya Pencegahan ISPA pada Balita di Puskesmas Ngoresan Surakarta . Skripsi. Jurusan Keperawatan Fakultas Ilmu Kesehatan. Universitas Muhammadiyah Surakarta.
- [13] Smeltzer. S. 2002. *Buku Ajar Keperawatan Medical Bedah*, volume 3, Jakarta: EGC.
- [14] Strehl, A., J. Ghosh, and R. Mooney. 2000. *Impact of Similarity Measures on Web-Page Clustering*, Proceeding of the Workshop of Artificial Intelligent for Web Search, 17th National Conference on Artificial Intelligence, July 2000.

- [15] Tan. Pang-Ning, Steinbach. Michael, Kumar. Vipin. 2006. *Introduction to Data Mining*, Pearson Education Inc.
- [16] Widjaja, Anton C, dr. 2003. Penanganan ISPA pada Anak di Rumah Sakit Kecil Negara Berkembang. Jakarta: Buku Kedokteran EGC.
- [17] World Health Organization. 2007. Pencegahan dan Pengendalian Infeksi Saluran Pernafasan Akut (ISPA) yang Cenderung Menjadi Pandemi dan Pandemi Fasilitas Pelayanan Kesehatan. Indonesia: Pedoman Interim WHO.