

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

Kanker merupakan pembelahan sel-sel tubuh yang tidak terkendali, penyakit yang mengancam jiwa wanita-wanita mulai dari tahap remaja hingga dewasa di seluruh dunia. Kanker serviks merupakan kasus yang menyerang organ intim wanita tepatnya di daerah leher rahim yaitu bagian paling bawah pada uterus dan hingga saat ini sudah ada ratusan ribu kasus kanker serviks dan puluhan ribu kasus kematian pada wanita mulai dari tahap usia remaja hingga tahap usia dewasa.[1] Kanker serviks adalah penyebab kanker terbesar nomor dua bagi wanita di seluruh dunia. Menurut data terkini dari WHO (Organisasi Kesehatan Dunia) setiap tahunnya 500,000 wanita diseluruh dunia didiagnosa mengidap kanker serviks dan sekitar 300,000 meninggal karena kanker serviks. Angka kematian pada wanita diakibatkan kanker serviks selalu meningkat dari tahun ke tahun. Tingkat kematian wanita dengan kanker serviks sangatlah tinggi dan hampir 60% diantaranya datang dari kaum-kaum yang menengah dan menengah kebawah dikarenakan individu tidak memiliki dana untuk melakukan pengecekan atau deteksi dini dan pengobatan untuk kanker serviks yang ada pada dalam tubuhnya.[2]

Data kanker global mengungkapkan bahwa kanker serviks adalah penyakit paling umum keempat di antara perempuan, dengan tingkat kematian sekitar 90% di negara-negara terbelakang dan berkembang karena tidak adanya pengetahuan - pengetahuan publik tentang penyebab dan dampaknya yang ditimbulkan dari kanker serviks. Penyebab utama kanker serviks adalah HPV (Human Papilloma Virus). Virus ini dapat membahayakan sel-sel serviks, sel skuamosa dan sel kelenjar. Sel-sel prakanker ini disebut sebagai Neoplasia Intraepithelial Serviks (CIN), yang mempengaruhi permukaan jaringan serviks saja. Seiring waktu, sebagian kecil CIN akan berkembang menjadi kanker. Untuk itu diperlukan periode dua hingga tiga dekade bagi kanker serviks untuk mencapai keadaan agresif, deteksi dini dan perawatan yang tepat dapat secara signifikan mengurangi penyakit ini.[3] Faktor-faktor yang terkait dengan perkembangan kanker serviks termasuk aktivitas seksual yang dimulai pada usia muda ( kurang dari 16 tahun), jumlah total pasangan seksual yang tinggi (lebih dari empat),

dan riwayat kutil kelamin. Sebenarnya, vaksin untuk virus seperti HPV 16 dan HPV 18 sekarang tersedia di pasaran, tetapi karena kurangnya kesadaran, pengetahuan dan juga sosialisasi yang baik, maka langkah-langkah pencegahan ini tidak umum digunakan. Beberapa tindakan pencegahan sederhana namun memiliki pengaruh besar yaitu kesadaran akan terjadinya kanker serviks, pernikahan dini atau di usia yang masih sangat belia, penggunaan pil kontrasepsi yang berkepanjangan, kebersihan organ intim, banyak pasangan seks atau bertukar-tukar pasangan, dan kekebalan atau daya tahan tubuh rendah. Namun, dari beberapa tindakan tersebut, hal paling penting yang harus dilakukan untuk mencegah terjadinya kanker serviks yaitu dengan melakukan deteksi dini dan perawatan yang tepat sehingga dapat mengurangi insiden dan resiko kematian yang tidak terdeteksi secara signifikan.[4]

Terdapat dua tes untuk mendeteksi perubahan awal kanker serviks, tes PAP Smear dan tes HPV DNA. Tes PAP (Papanicolaou test) adalah tes pemindaian yang lebih baik dan umum digunakan untuk deteksi dini kanker serviks.[5] Tes PAP Smear pertama kali ditunjukkan oleh ilmuwan George Papanicolaou pada tahun 1940 untuk membantu dalam mendeteksi perubahan prekanker pada sel-sel serviks. Dalam tes PAP Smear, sel-sel dihapus dari serviks dan kemudian disebarkan di atas kaca slide. Lalu sel sel kemudian diwarnai dengan pewarna dan dibiarkan kering. Dalam melakukan pemindaian untuk kanker serviks juga membutuhkan sejumlah biaya dimana kaum wanita yang menengah dan menengah kebawah akan sangat dirugikan karena tidak memiliki dana yang mencukupi untuk melakukan pengecekan sedangkan bagi kaum yang menengah ke atas akan bisa selalu melakukan pemindaian serta pengecekan dan memproteksi diri seperti masuk ke dalam asuransi kesehatan. Begitu juga dalam tes HPV DNA yang mendeteksi adanya unsur genetik (DNA) dari virus HPV yang beresiko tinggi menimbulkan kanker serviks. Misalnya, untuk wanita yang mengalami kutil kelamin, maka dokter spesialis kanker akan menganjurkan penderita untuk melakukan tes kembali dalam waktu satu tahun, bertujuan untuk mengetahui bahwa penderita masih terinfeksi virus HPV serta mengetahui perubahan sel-sel pada serviks(leher rahim).[6]

Machine Learning telah digunakan dalam deteksi dan diagnosis kanker selama hampir 20 tahun. metode Machine Learning sedang digunakan dalam berbagai aplikasi mulai dari mendeteksi dan mengklasifikasikan tumor melalui X-ray dan gambar CRT

hingga klasifikasi keganasan dari proteomik dan genomik. Para peneliti kanker telah mencoba menerapkan Machine Learning terhadap prediksi dan prognosis kanker dalam mengidentifikasi, mengklasifikasikan, mendeteksi, atau membedakan tumor dan keganasan lainnya.[7] Pemindaian sel kanker secara manual sering menghasilkan variasi besar dalam kualitas spesimen, seperti distribusi bahan seluler yang tidak rata yang dapat menyebabkan rumpun padat cahaya yang tidak dapat menembus sedangkan bagian lain dari spesimen mungkin memiliki banyak sel yang tumpang tindih yang menghambat interpretasi yang akurat. Selain itu, pemeriksaan visual manual memakan waktu dan analisis serta klasifikasi ratusan atau ribuan sel bisa tidak akurat karena kesalahan manusia. Ketika pemeriksaan sel untuk kelainan dilakukan oleh komputer, sel harus dipindai dengan resolusi tinggi untuk mengekstraksi fitur dengan andal. Karena variasi ukuran dan bentuk sel normal dan abnormal, segmentasi dan klasifikasi sel yang akurat sangat penting untuk membedakan antara sel normal dan abnormal.[8] Saat ini ketergantungan domain medis pada sistem yang dibantu komputer telah meningkat. Ada banyak sistem otomatis yang mewakili gambar digital menggunakan teknik Machine Learning. Ekstraksi fitur harus cukup kuat untuk mewakili gambar dalam sistem digital. Dalam pengenalan pola dan klasifikasi Machine Learning, fitur gambar memainkan peran penting untuk analisis gambar. Dalam sistem analisis citra medis, ekstraksi fitur informatif dari gambar adalah tugas paling penting karena rendahnya iluminasi gambar medis. Banyak sistem otomatis dan semi-otomatis telah diusulkan dalam sepuluh tahun terakhir, beberapa bekerja pada satu sel, dan beberapa bekerja pada banyak sel.[9]

## **1.2. Rumusan Masalah**

Adapun rumusan masalah penelitian ini yaitu: Bagaimana analisa perbandingan metode Random Forest Tree dan K-Nearest Neighbor dalam melakukan pendeteksian Kanker Serviks?

### **1.3. Tujuan dan Manfaat**

#### **1.3.1. Tujuan**

Adapun tujuan dari penelitian ini adalah untuk mengetahui analisa perbandingan antara metode Random Forest Tree dan K-Nearest Neighbor dalam melakukan pendeteksian Kanker Serviks.

#### **1.3.2. Manfaat**

Manfaat dari penelitian ini diantaranya adalah:

1. Membantu pendeteksian Kanker Serviks menggunakan metode Random Forest Tree dan K-Nearest Neighbor.
2. Membantu dokter dalam mendeteksi Kanker Serviks dengan cepat.
3. Mengetahui metode yang lebih cocok dalam mengklasifikasi Kanker Serviks.

### **1.4. Batasan Masalah**

Dalam penelitian ini, terdapat batasan masalah dari permasalahan yang ada, yaitu :

1. Dataset yang digunakan adalah dataset dari yang direkam pada rumah sakit Hospital Universitario de Caracas di Caracas, Venezuela[10].
2. Jenis metode yang digunakan adalah metode Random Forest Tree dan K-Nearest Neighbor yaitu pendeteksian dini bagi yang penderita Kanker Serviks.
3. Bahasa pemrograman yang digunakan adalah bahasa Python dibantu dengan library Keras.
4. Cakupan perbandingan adalah dalam hal akurasi, dan kecepatan training.

### **1.5. Keterbaruan**

Random Forest Tree adalah algoritma klasifikasi klasik yang telah sering dipakai dan diteliti. Salah satu penelitian terakhir mengenai algoritma Random Forest dilakukan oleh W. Chen dalam penelitian yang berjudul Groundwater Spring Potential Mapping Using Artificial Intelligence Approach Based on Kernel Logistic Regression, Random

Forest, and Alternating Decision Tree Models. Random Forest dapat dipakai untuk menganalisa dataset secara regresi maupun klasifikasi. Dalam penelitiannya, performa Kernel Logistic Regression, Random Forest dan Alternating Decision Tree Models dinilai dengan nilai Area Under Curve (AUC) dengan nilai 0.909, 0.8777 dan 0.812. Akurasi adalah proporsi dari true positives dan negatives dari keseluruhan dataset sedangkan AUC mengukur selisih antara true positive rate (recall) dan false positive rate[11].

KNN juga sebelumnya telah digunakan untuk menganalisa dataset COVID-19 pada penelitian yang dilakukan oleh W.M. Shaban dengan judul A new COVID-19 Patients Detection Strategy (CPDS) based on hybrid feature selection and enhanced KNN Classifier. Penelitian tersebut berfokus pada penemuan strategi diagnosa yang diberi nama COVID-19 Patients Detection Strategy (CPDS). Penelitian ini menawarkan dua kontribusi baru yaitu: merupakan metode fitur seleksi hybrid(HFSM) dan menggunakan algoritma klasifikasi enhanced K-Nearest Neighbor (EKNN) dimana algoritma KNN ditambahkan heuristik solid dalam pemilihan neighbors. Hasil penelitian ini menunjukkan akurasi maksimum dari metode strategi deteksi yang diusulkan[12].

Penelitian terakhir mengenai klasifikasi kanker serviks dilakukan oleh Z. Alyafeai dengan jurnal yang berjudul “A fully-automated deep learning pipeline for cervical cancer classification”. Pada penelitian tersebut beberapa dataset diuji dengan model yang dibangun menggunakan algoritma Convolutional Neural Network (CNN). Lima model yang dibangun dengan menggunakan lima dataset diuji dan diukur berdasarkan tingkat akurasi, spesififikasi dan sensitivitas[13]. Peneliti dalam penelitian ini membandingkan performa algoritma Random Forest Tree dengan KNN terhadap database kanker serviks untuk menganalisa kedua algoritma tersebut dan dampaknya terhadap model yang dibangun untuk mendeteksi kanker serviks.