

BAB II

KERANGKA TEORITIS

A. Landasan Teori

1. Pengertian Data Mining

Mengutip (Tan, 2006 dalam Eko Prasetyo, 2013:2) menyatakan bahwa data mining adalah sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Data mining juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah data mining disebut juga knowledge discovery. Knowledge Discovery in Databases (KDD) adalah penerapan metode saintifik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD (Turban, dkk 2001). KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah data. Serangkaian proses tersebut memiliki tahap sebagai berikut (Tan, 2006) :

- a. Pembersihan data yaitu untuk membuang data yang tidak konsisten dan noise
- b. Integrasi data yaitu penggabungan data dari beberapa sumber
- c. Transformasi data yaitu data diubah menjadi bentuk yang sesuai untuk di mining
- d. Aplikasi teknik data mining yaitu proses ekstraksi pola dari data yang ada
- e. Evaluasi pola yang ditemukan yaitu proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan
- f. Presentasi pengetahuan yaitu dengan teknik visualisasi

Tahap ini merupakan bagian dari proses pencarian pengetahuan yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya. Langkah terakhir KDD adalah mempresentasikan pengetahuan dalam bentuk yang mudah dipahami pengguna (Tan, 2006).

2. Pengertian Klasifikasi

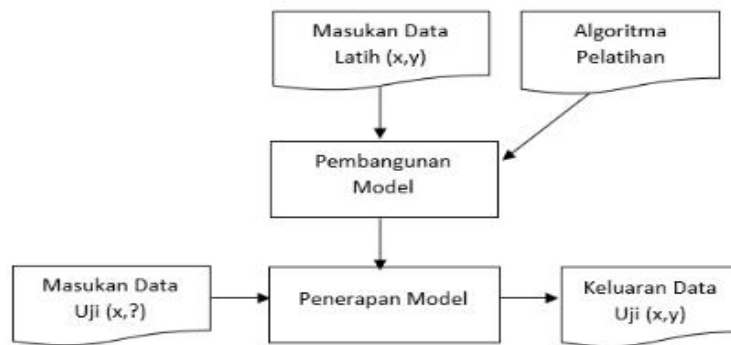
Klasifikasi merupakan suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang ada yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Aturan-aturan tersebut digunakan pada data-data baru untuk diklasifikasi.

Teknik ini menggunakan supervised induction, yang memanfaatkan kumpulan pengujian dari record yang terklasifikasi untuk menentukan kelas-kelas tambahan (Kusnawi, 2007). Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu (1) pembangunan model

sebagai prototipe untuk disimpan sebagai memori dan (2) penggunaan model tersebut untuk melakukan pengenalan/klasifikasi/prediksi pada suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang sudah disimpannya (Eko Prasetyo, 2012).

Klasifikasi dapat didefinisikan secara detail sebagai waktu pekerjaan yang melakukan pelatihan atau pembelajaran terhadap fungsi target / yang memetakan setiap vektor (set fitur) x ke depan satu dari sejumlah tabel kelas y yang tersedia pekerjaan pelatihan tersebut akan menghasilkan suatu model yang kemudian disimpan sebagai memori (Eko Prasetyo, 2013).

Model dalam dalam klasifikasi mempunyai arti yang sama dengan kotak hitam, dimana ada suatu model yang menerima masukan, kemudian mampu melakukan pemikiran terhadap masukan tersebut dan memberikan jawaban sebagai keluaran dari hasil pemikirannya. Kerangka kerja (*framework*) klasifikasi ditunjukkan pada Gambar 2.1. Pada gambar tersebut disediakan sejumlah data latih (x,y) untuk digunakan sebagai data pembangun model. Model tersebut kemudian dipakai untuk memprediksi kelas dari data uji ($x,?$) sehingga diketahui kelas y yang sesungguhnya.



Gambar 2. 1 Proses Pekerjaan Klasifikasi

Klasifikasi digunakan untuk pembuatan model yang dapat melakukan pemetaan dari setiap himpunan variabel kesetiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan variabel yang baru didapat. Dalam pembangunan model selama proses pelatihan tersebut diperlukan adanya suatu algoritma untuk membangunnya yang disebut dengan Algoritma pelatihan. Ada banyak algoritma pelatihan yang sudah dikembangkan oleh para peneliti yaitu C4.5, Nearest Neighbor, Bayesian Classification, Neural Network, dll.

3. Pengembangan Sistem SDLC

Pendekatan sistem merupakan sebuah metodologi. Metodologi adalah satu cara yang direkomendasikan dalam melakukan sesuatu. Pendekatan sistem adalah metodologi dasar dalam memecahkan segala jenis masalah. Siklus hidup pengembangan sistem (System Development Life Cycle – SDLC) adalah aplikasi dari pendekatan sistem bagi pengembangan suatu sistem informasi (Raymond McLeod 2007: 199).

Terdapat beberapa tahapan pekerjaan pengembangan yang perlu dilakukan jika suatu proyek ingin memiliki kemungkinan berhasil yang besar. Tahapan-tahapan tersebut adalah:

- a. Perencanaan
- b. Analisis
- c. Desain
- d. Implementasi
- e. Penggunaan

Proyek dan sumber daya yang dibutuhkan untuk melakukan pekerjaan direncanakan kemudian disatukan. Sistem yang ada juga dianalisis untuk memahami masalah dan menentukan persyaratan fungsional dari sistem yang baru. Sistem baru ini kemudian dirancang dan diimplementasikan. Setelah implementasi, sistem kemudian digunakan, idealnya untuk jangka waktu yang lama.

Karena pekerjaan-pekerjaan di atas mengikuti satu pola yang teratur dan dilaksanakan dengan cara dari atas ke bawah, SDLC tradisional sering kali disebut sebagai pendekatan air terjun (waterfall approach). Aktivitas ini memiliki aliran satu arah menuju ke penyelesaian proyek.



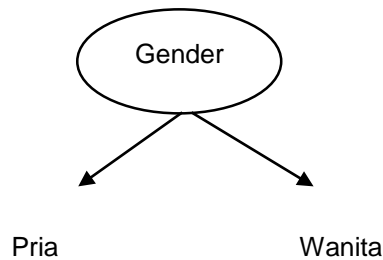
Gambar 2.1 Pola Melingkar dari Siklus Hidup Sistem
(Sumber: Raymond McLeod 2007: 199)

Gambar 2.1 mengilustrasikan sifat melingkar dari siklus hidup. Ketika sebuah sistem telah melampaui masa manfaatnya dan harus diganti, satu siklus hidup baru akan dimulai dengan diawali oleh tahap perencanaan.

Mudah bagi kita untuk melihat bagaimana SDLC tradisional dapat dikatakan sebagai suatu aplikasi dari pendekatan sistem. Masalah akan didefinisikan dalam tahap-tahapan perencanaan dan analisis. Solusi-solusi alternatif diidentifikasi dan dievaluasi dalam tahap desain. Lalu, solusi yang terbaik diimplementasikan dan digunakan. Selama tahap penggunaan, umpan balik dikumpulkan untuk melihat seberapa baik sistem mampu memecahkan masalah yang telah ditentukan.

4. Metode Algoritma C4.5

Algoritma C4.5 adalah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari record yang baru. Algoritma C4.5 sendiri merupakan pengembangan dari algoritma ID3, dimana pengembangan dilakukan dalam hal, bisa mengatasi *missing data*, bisa mengatasi data kontinu dan *pruning* (Larose, 2005).



Gambar 2. 2 Syarat pengujian fitur biner

Yang menjadi hal penting dalam induksi decision tree adalah bagaimana menyatakan syarat pengujian pada node. Ada 3 kelompok penting dalam syarat pengujian node:

a. Fitur biner

Fitur yang hanya mempunyai dua nilai berbeda disebut dengan fitur biner. Syarat pengujian ketika fitur ini menjadi node (akar maupun internal) hanya punya dua pilihan cabang. Contoh pemecahannya disajikan pada Gambar 2.1.

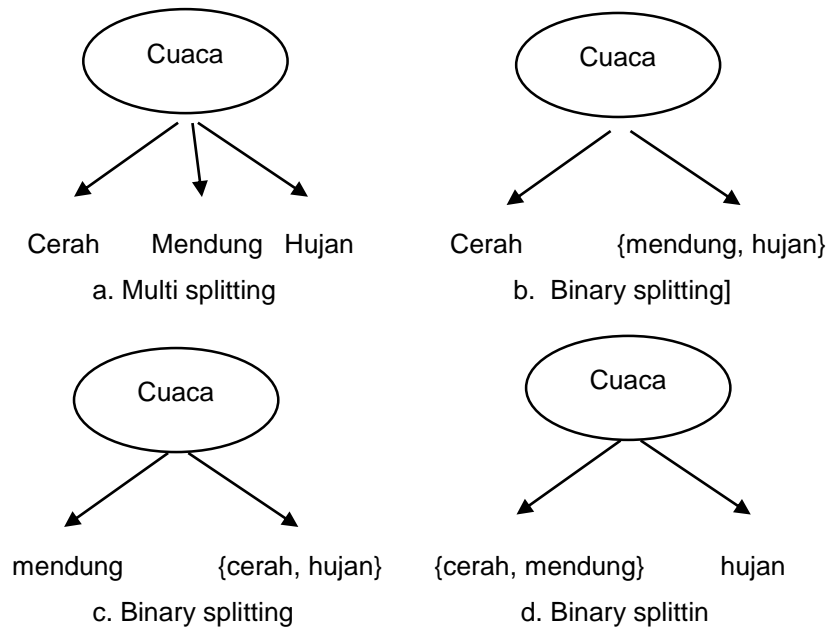
b. Fitur bertipe kategorikal

Untuk fitur yang nilainya bertipe kategorikal (nominal atau ordinal) bisa mempunyai beberapa nilai berbeda. Contohnya adalah fitur 'cuaca' mempunyai 3 nilai berbeda dan ini bisa mempunyai banyak kombinasi syarat pengujian pemecahan. Secara umum ada 2, yaitu pemecahan biner (*binary splitting*) dan (*multi splitting*). Kombinasinya disajikan seperti pada Gambar 2.2. Untuk pemecahan yang hanya membolehkan pemecahan biner, seperti algoritma CART, maka akan memberikan kemungkinan jumlah kombinasi pemecahan sebanyak 2^{k-1} , dimana k adalah jumlah nilai berbeda dalam fitur tersebut. Contoh pemecahannya disajikan pada Gambar 2.2 dan Gambar 2.3.

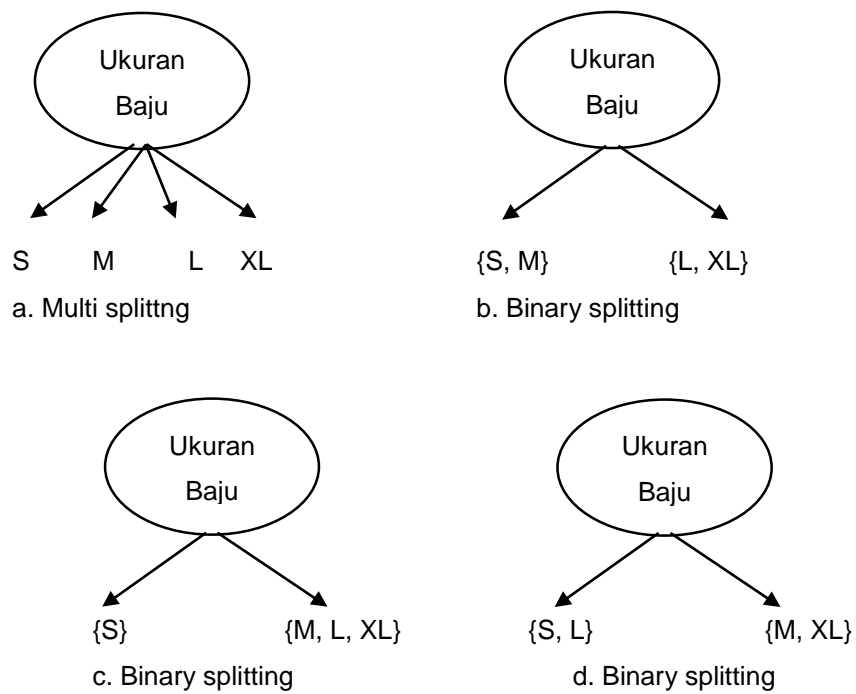
c. Fitur bertipe numerik

Untuk fitur bertipe numerik, syarat pengujian dalam node (akar maupun internal) dinyatakan dengan pengujian perbandingan ($A < v$) atau ($A \geq v$) dengan hasil biner, atau untuk multi dengan hasil berupa jangkauan nilai dalam bentuk $v_i \leq A < v_{i+1}$, untuk $i = 1, 2, \dots, k$. Untuk kasus pemecahan biner, maka algoritma akan memeriksa semua kemungkinan posisi pemecahan v dan memilih posisi v terbaik. Untuk cara multi, maka algoritma harus memeriksa semua kemungkinan

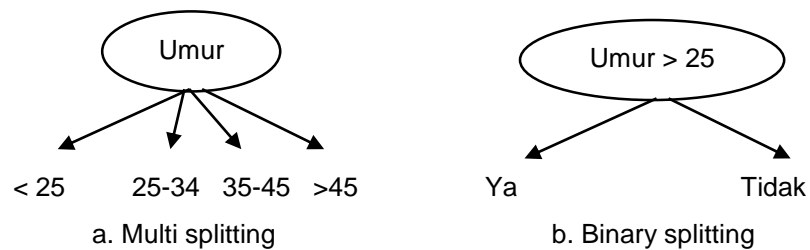
jangkauan nilai *kontinyu*. Contoh pemecahan pada fitur numerik disajikan pada gambar 2.4.



Gambar 2. 3 Syarat pengujian fitur bertipe nominal



Gambar 2. 4 Syarat pengujian fitur bertipe ordinal



Gambar 2. 5 Syarat pengujian fitur bertipe numerik

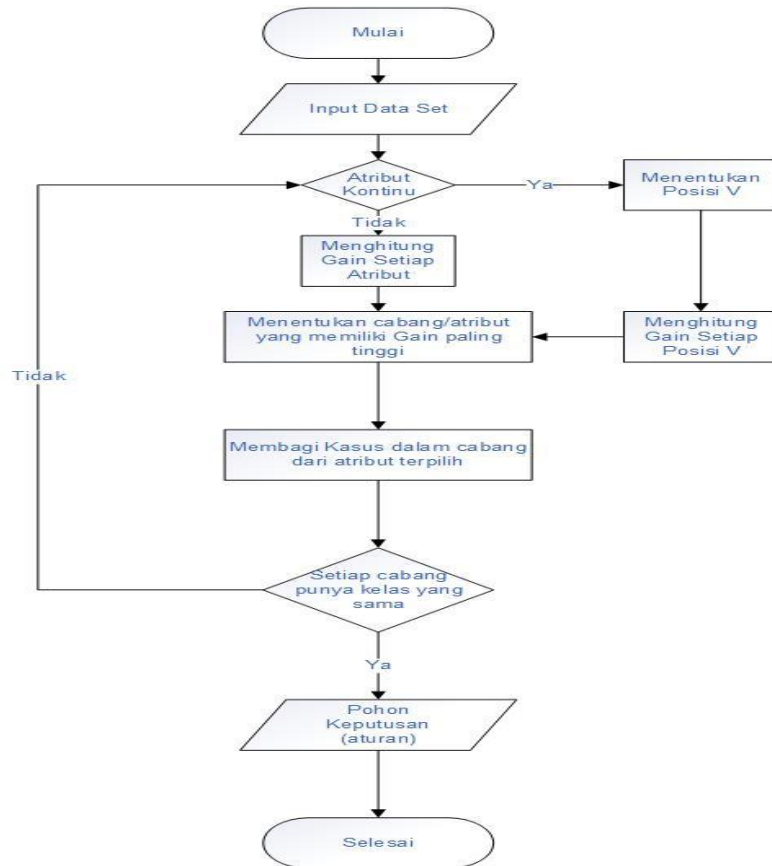
Kriteria yang paling banyak digunakan untuk memilih fitur sebagai pemecah dalam algoritma C4.5 adalah rasio gain, yang diformulasikan oleh persamaan berikut:

$$RasioGain(s,j) = \frac{Gain(s,j)}{SplitInfo(s,j)}$$

Persamaan diatas menyatakan nilai rasio gain pada fitur ke- j . $SplitInfo(s,j)$ didapat dari $SplitInfo(s,j) = -\sum_{i=1}^k p(v_i|s) \log_2 p(v_i|s)$ dimana k menyatakan jumlah pemecahan .

Berikut ini adalah algoritma C4.5 menggunakan flowcart yang disajikan pada Gambar 2.6 dan berserta penjelasannya :

1. Memasukan Data Set yang telah disediakan.Data Set adalah kumpulan objek dan atributnya.
2. Pada proses cek keputusan dilihat apakah data set tersebut termasuk Atribut Kontinu atau bukan, Atribut Kontinu adalah atribut yang mempunyai jangkauan real.
3. Jika Iya maka akan menentukan posisi V , setelah itu menghitung gain setiap V .Jika tidak maka menghitung gain setiap atribut.
4. Setelah itu menentukan cabang/atribut yang memiliki gain paling tinggi.
5. Setelah itu membagi kasus dalam cabang dari atribut yang terpilih.
6. Jika setiap cabang mempunyai kelas yang sama, maka menghasilkan pohon keputusan. Jika tidak maka cabang atau atribut tersebut kembali melakukan cek atribut kontinu.



Gambar 2. 4 Flowchart Algoritma C4.5

Sumber : (Larose, 2005)

5. PHP

PHP atau kependekan dari Hypertext Preprocessor adalah salah satu bahasa pemrograman open source yang sangat cocok atau dikhususkan untuk pengembangan web dan dapat ditanamkan pada sebuah skripsi HTML. Bahasa PHP dapat dikatakan menggambarkan beberapa bahasa pemrograman seperti C, Java, dan Perl serta mudah untuk dipelajari. PHP merupakan bahasa scripting server – side, dimana pemrosesan datanya dilakukan pada sisi server. Sederhananya, serverlah yang akan menerjemahkan skrip program, baru kemudian hasilnya akan dikirim kepada client yang melakukan permintaan. Adapun pengertian lain PHP adalah akronim dari Hypertext Preprocessor, yaitu suatu bahasa pemrograman berbasis kode – kode (script) yang digunakan untuk mengolah suatu data dan mengirimkannya kembali ke web browser menjadi kode HTML”. Menurut Kustiyarningsih (2011:114), “PHP (atau resminya PHP: Hypertext Preprocessor) adalah skrip bersifat server – side yang ditambahkan ke dalam HTML”. Teoritis terkait dengan objek permasalahan

B. Teoritis terkait dengan objek permasalahan

1. Pengertian Tanah Longsor

Gerakan tanah adalah suatu gerakan menuruni lereng oleh massa tanah dan atau batuan penyusun lereng. Definisi di atas dapat menunjukkan bahwa massa yang bergerak dapat berupa massa tanah, massa batuan ataupun percampuran antara keduanya. Masyarakat pada umumnya menerapkan istilah longsor untuk seluruh jenis gerakan tanah, baik yang melalui bidang gelincir ataupun tidak. Varnes (1978) secara definitif juga menerapkan istilah longsor ini untuk seluruh jenis gerakan tanah. Gerakan tanah merupakan salah satu proses geologi yang terjadi akibat interaksi beberapa kondisi antara lain geomorfologi, struktur geologi, hidrogeologi dan tata guna lahan. Kondisi tersebut saling berpengaruh sehingga mewujudkan kondisi lereng yang cenderung bergerak (Karnawati, 2007). Gerakan tanah dapat diidentifikasi melalui tanda-tanda sebagai berikut: munculnya retak tarik dan kerutan-kerutan di permukaan lereng, patahnya pipa dan tiang listrik, miringnya pepohonan, perkerasan jalan yang terletak pada timbunan mengalami amblas, rusaknya perlengkapan jalan seperti pagar pengaman dan saluran drainase, tertutupnya sambungan ekspansi pada pelat jembatan, hilangnya kelurusan dari fondasi bangunan, tembok bangunan retakretak, dan dinding penahan tanah retak serta miring ke depan (Hardiyatmo, 2012). 9 Kerentanan lereng terhadap gerakan tanah didefinisikan sebagai kecenderungan lereng dalam suatu wilayah atau zona untuk mengalami gerakan, tanpa mempertimbangkan resikonya terhadap kerugian jiwa atau ekonomi. Apabila aspek risiko terhadap manusia diperhitungkan, maka lebih tepat diterapkan istilah kerawanan (BAPEKOINDA, 2002).

2. Jenis Tanah Longsor

Varnes (1978) mengklasifikasi tanah longsor menjadi 6 jenis yaitu runtuh (*fall*), robohan (*topple*), longsor (*slides*), pencaran lateral (*lateral spread*), aliran (*flow*) dan gabungan. Klasifikasi Varnes didasarkan pada mekanisme gerakan dan material yang berpindah. Klasifikasi tersebut diuraikan sebagai berikut:

Runtuhan (*falls*) adalah runtuhnya sebagian massa batuan pada lereng yang terjal, seperti ditunjukkan pada Gambar 2. Jenis ini memiliki ciri yaitu sedikit atau tanpa disertai terjadinya pergeseran antara massa yang runtuh dengan massa yang tidak runtuh. Runtuhnya massa batuan umumnya dengan cara jatuh bebas, meloncat atau menggelinding tanpa melalui bidang gelincir. Penyebab terjadinya runtuh adalah adanya bidang-bidang diskontinyu seperti retakan- retakan pada batuan.



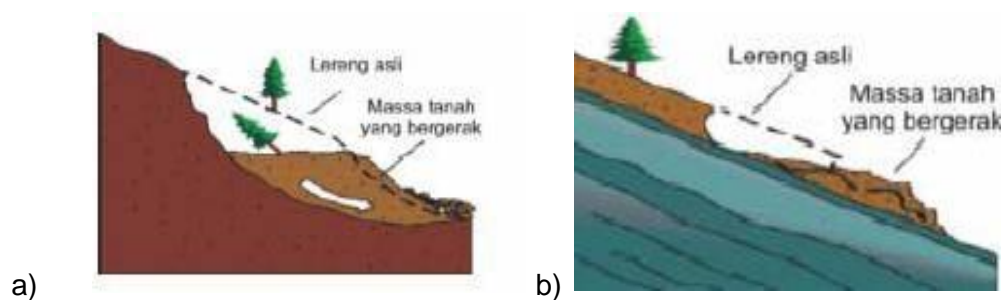
**Gambar 2. 5 Runtuhan Batuan
(Rahmawati, 2009)**

Robohan (*topples*) adalah robohnya batuan umumnya bergerak melalui bidang-bidang diskontinyu yang sangat tegak pada lereng. Bidang diskontinyu ini berupa retakan pada batuan seperti pada runtuh. Robohan ini biasanya terjadi pada batuan dengan kelerengan sangat terjal sampai tegak, seperti ditunjukkan pada Gambar 3.



**Gambar 2. 6 Robohan Batuan
(Rahmawati, 2009)**

Longsor (*Slide*) adalah gerakan menuruni lereng oleh material penyusun lereng, melalui bidang gelincir pada lereng. Seringkali dijumpai tanda-tanda awal gerakan berupa retakan berbentuk lengkung tapal kuda pada bagian permukaan lereng yang mulai bergerak. Bidang gelincir ini dapat berupa bidang yang relatif lurus (translasi) ataupun bidang lengkung ke atas (rotasi), seperti ditunjukkan pada Gambar 2.9



Gambar 2. 7 . (a) Rotasi Batuan (b) Luncuran batuan

(Rahmawati, 2009)

Kedalaman bidang gelincir pada longsoran jenis translasi lebih dangkal daripada kedalaman bidang gelincir longsoran rotasi. Material yang bergerak secara translasi dapat berupa blok (*rock block slide*). Longsoran yang bergerak secara rotasi melalui bidang gelincir lengkung disebut nendatan (*slump*). Nendatan umumnya terjadi pada lereng yang tersusun oleh material yang relatif homogen.

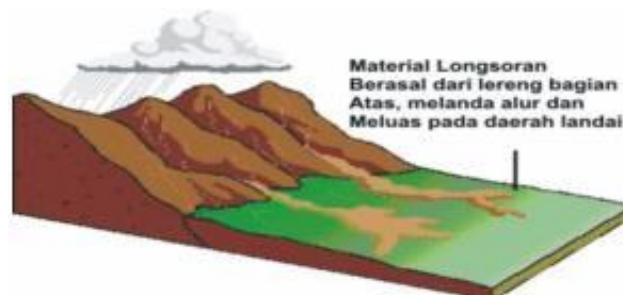
Pencaran lateral (*lateral spread*) adalah material tanah atau batuan yang bergerak dengan cara perpindahan translasi pada bagian dengan kemiringan landai sampai datar. Pergerakan terjadi pada lereng yang tersusun atas tanah lunak dan terbebani oleh massa tanah di atasnya, seperti ditunjukkan pada Gambar 5. Pembebanan inilah yang mengakibatkan lapisan tanah lunak tertekan dan mengembang ke arah lateral.



Gambar 2. 8 Pencaran Batuan

(Rahmawati, 2009)

Aliran (*flows*) yaitu aliran massa yang berupa aliran fluida kental, seperti ditunjukkan pada Gambar 6. Aliran pada bahan rombakan dapat dibedakan menjadi aliran bahan rombakan (*debris*), aliran tanah (*earth flow*) apabila massa yang bergerak didominasi oleh material tanah berukuran butir halus (butir lempung) dan aliran lumpur (*mud flow*) apabila massa yang bergerak jenuh air. Jenis lain dari aliran ini adalah aliran kering yang biasa terjadi pada endapan pasir (*dry flow*).



Gambar 2. 9 Aliran Batuan

(Rahmawati, 2009)

Di alam sering terjadi tanah longsor dengan mekanisme gabungan dari dua atau lebih jenis tanah longsor. Tanah longsor tersebut diklasifikasikan sebagai tanah longsor gabungan atau kompleks.

3. Gejala Umum Tanah Longsor

- a. Munculnya Retakan – retakan di lereng yang sejajar dengan arah tebing.
- b. Biasanya terjadi setelah hujan.
- c. Munculnya mata air baru setelah tiba-tiba.
- d. Tebing rapuh dan kerikil berjatuhan.

4. Penyebab terjadinya Tanah Longsor

Pada prinsipnya tanah longsor terjadi bila gaya pendorong pada lereng lebih besar daripada gaya penahan. Gaya penahan umumnya dipengaruhi oleh kekuatan batuan dan kepadatan tanah. Sedangkan gaya pendorong dipengaruhi oleh besarnya sudut lereng, air, beban serta berat jenis tanah batuan.

5. Faktor-faktor Penyebab Tanah Longsor

a. Hujan

Ancaman tanah longsor biasanya dimulai pada bulan November karena meningkatnya intensitas curah hujan. Musim kering yang panjang akan menyebabkan terjadinya penguapan air di permukaan tanah dalam jumlah besar. Hal itu mengakibatkan munculnya pori-pori atau rongga tanah hingga terjadi retakan dan merekahnya tanah permukaan. Ketika hujan, air akan menyusup ke bagian yang retak sehingga tanah dengan cepat mengembang kembali. Pada awal musim hujan, intensitas hujan yang tinggi biasanya sering terjadi, sehingga kandungan air pada tanah menjadi jenuh dalam waktu singkat. Hujan lebat pada awal musim dapat menimbulkan longsor, karena melalui tanah yang merekah air akan masuk dan terakumulasi di bagian dasar lereng, sehingga menimbulkan gerakan lateral. Bila ada pepohonan di permukaannya, tanah longsor dapat dicegah karena air akan diserap oleh tumbuhan. Akar tumbuhan juga akan berfungsi mengikat tanah.

b. Lereng Terjal

Lereng atau tebing yang terjal akan memperbesar gaya pendorong. Lereng yang terjal terbentuk karena pengikisan air sungai, mata air, air laut, dan angin. Kebanyakan sudut lereng yang menyebabkan longsor adalah 180 apabila ujung lerengnya terjal dan bidang longsorannya mendatar.

c. Tanah yang kurang padat dan tebal

Jenis tanah yang kurang padat adalah tanah lempung atau tanah liat dengan ketebalan lebih dari 2,5 m dan sudut lereng lebih dari 220. Tanah jenis ini memiliki potensi untuk terjadinya tanah longsor terutama bila terjadi hujan. Selain itu tanah ini sangat rentan terhadap pergerakan tanah karena menjadi lembek terkena air dan pecah ketika hawa terlalu panas.

d. Batuan yang kurang kuat

Batuan endapan gunung api dan batuan sedimen berukuran pasir dan campuran antara kerikil, pasir, dan lempung umumnya kurang kuat. Batuan tersebut akan mudah menjadi tanah bila mengalami proses pelapukan dan umumnya rentan terhadap tanah longsor bila terdapat pada lereng yang terjal.

e. Jenis tata lahan

Tanah longsor banyak terjadi di daerah tata lahan persawahan, perladangan, dan adanya genangan air di lereng yang terjal. Pada lahan persawahan akarnya kurang kuat untuk mengikat butir tanah dan membuat tanah menjadi lembek dan jenuh dengan air sehingga mudah terjadi longsor. Sedangkan untuk daerah perladangan penyebabnya adalah karena akar pohonnya tidak dapat menembus bidang longsor yang dalam dan umumnya terjadi di daerah longsor lama.

f. Getaran

Getaran yang terjadi biasanya diakibatkan oleh gempa bumi, ledakan, getaran mesin, dan getaran lalu lintas kendaraan. Akibat yang ditimbulkannya adalah tanah, badan jalan, lantai, dan dinding rumah menjadi retak.

g. Susut muka air danau atau bendungan

Akibat susutnya muka air yang cepat di danau maka gaya penahan lereng menjadi hilang, dengan sudut kemiringan waduk 220 mudah terjadi longsor dan penurunan tanah yang biasanya diikuti oleh retakan.

h. Adanya beban tambahan

Adanya beban tambahan seperti beban bangunan pada lereng, dan kendaraan akan memperbesar gaya pendorong terjadinya longsor, terutama di sekitar tikungan jalan pada daerah lembah. Akibatnya adalah sering terjadinya penurunan tanah dan retakan yang arahnya ke arah lembah.

i. Pengikisan/erosi

Pengikisan banyak dilakukan oleh air sungai ke arah tebing. Selain itu akibat penggundulan hutan di sekitar tikungan sungai, tebing akan menjadi terjal

j. Adanya material timbunan pada tebing

Untuk mengembangkan dan memperluas lahan pemukiman umumnya dilakukan pemotongan tebing dan penimbunan lembah. Tanah timbunan pada lembah tersebut belum terpadatkan sempurna seperti tanah asli yang berada di bawahnya. Sehingga apabila hujan akan terjadi penurunan tanah yang kemudian diikuti dengan retakan tanah.

k. Bekas longsor lama

Longsor lama umumnya terjadi selama dan setelah terjadi pengendapan material gunung api pada lereng yang relatif terjal atau pada saat atau sesudah terjadi patahan kulit bumi. Bekas longsor lama memiliki ciri:

- 1) Adanya tebing terjal yang panjang melengkung membentuk tapal kuda.
- 2) Umumnya dijumpai mata air, pepohonan yang relatif tebal karena tanahnya

gembur dan subur.

- 3) Daerah badan longsor bagian atas umumnya relatif landai.
 - 4) Dijumpai longsor kecil terutama pada tebing lembah.
 - 5) Dijumpai tebing-tebing relatif terjal yang merupakan bekas longsor kecil pada longsor lama.
 - 6) Dijumpai alur lembah dan pada tebingnya dijumpai retakan dan longsor kecil.
 - 7) Longsor lama ini cukup luas.
- l. Adanya bidang diskontinuitas (bidang tidak sinambung) bidang tidak sinambung memiliki ciri :
- 1) Bidang perlapisan batuan
 - 2) Bidang kontak antara tanah penutup dengan batuan dasar
 - 3) Bidang kontak antara batuan yang retak-retak dengan batuan yang kuat.
 - 4) Bidang kontak antara batuan yang dapat melewatkan air dengan batuan yang tidak melewatkan air (kedap air).
 - 5) Bidang kontak antara tanah yang lembek dengan yang padat.
 - 6) Bidang-bidang tersebut merupakan bidang lemah dan dapat berfungsi sebagai bidang luncuran tanah longsor.
- m. Penggundulan hutan
Tanah longsor umumnya banyak terjadi di daerah yang relatif gundul dimana pengikatan air tanah sangat kurang.
- n. Daerah pembuangan sampah
Penggunaan lapisan tanah yang rendah untuk pembuangan sampah dalam jumlah banyak dapat mengakibatkan tanah longsor apalagi ditambah dengan guyuran hujan, seperti yang terjadi di Tempat Pembuangan Akhir Sampah Leuwigajah di Cimahi. Bencana ini menyebabkan sekitar 120 orang lebih meninggal.

C. Tinjauan Studi

Penelitian rujukan merupakan acuan yang dibutuhkan seorang peneliti untuk melakukan penelitian. Penelitian rujukan pada penelitian ini diambil berdasarkan kesamaan metode yang digunakan yaitu Algoritma C4.5. Banyak penelitian yang menggunakan metode ini dalam berbagai kasus. Antara lain :

1. **Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada (Muhammad Fauzul Arifin, Devi Fitriyah 2018) (Arifin, Fitriyah 2018).** Permasalahan yang sering muncul dalam bisnis pada penjualan dengan sistem pembayaran kredit tempo adalah antara lain seperti kredit macet, order fiktif dan penipuan. Jika tidak menggunakan prosedur yang benar dalam penerimaan mitra penjualan bukan tidak mungkin masalah tersebut akan timbul dalam proses bisnis. Untuk mengembangkan bisnis tentu perlu meningkatkan penjualan dan dibarengi dengan pembayaran yang lancar. Oleh karena itu, pada proses

penerimaan mitra perlu dibuatkan prosedur yang baik agar dapat meminimalisir kemungkinan masalah yang ada. Pada prosedur penerimaan mitra penjualan perlu dilakukan proses analisa untuk menentukan calon mitra penjualan tersebut diterima atau ditolak. Dari data mitra yang ada dapat digunakan untuk dilakukan pengolahan data. Salah satu teknik pengolahan data yang dapat digunakan dalam proses tersebut adalah klasifikasi. Klasifikasi adalah teknik pengolahan data yang membagi objek menjadi beberapa kelas sesuai dengan jumlah kelas yang diinginkan. Dan menggunakan Algoritma C4.5 dalam proses pengklasifikasi-annya. Yang dapat menentukan apakah calon mitra dapat diterima sebagai mitra atau tidak. Kemudian, Hasil dari pengklasifikasiannya divalidasi dengan ten-fold cross validation dengan tingkat akurasi 96,26 %, presisi 100% dan recall 71,43%.

2. Penerapan Algoritma C4.5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE (Fandy Ferdian Harryanto, Seng Hansun, 2017) (Ferdian,Hansun, 2017)

Perusahaan pada umumnya memerlukan pegawai yang memiliki kemampuan yang baik, perilaku yang baik serta dapat menyelesaikan pekerjaan yang diberikan kepadanya. Namun terdapat beberapa kesulitan untuk mengetahui kualitas orang-orang yang memiliki potensi baik sebagai pegawai pada suatu perusahaan. Oleh karena itu diperlukan cara atau metode untuk mengidentifikasi calon pegawai suatu perusahaan. Algoritma C4.5 dapat digunakan untuk melakukan prediksi dan klasifikasi terhadap calon pegawai yang berpotensi untuk masuk ke dalam perusahaan dengan cara membuat pohon keputusan berdasarkan data yang sudah ada dan melakukan prediksi terhadap calon pegawai baru yang ingin masuk ke perusahaan. Berdasarkan metode pengukuran akurasi ten-fold cross validation telah didapatkan hasil pengukuran tingkat keberhasilan prediksi calon pegawai baru sebesar 71% dengan menggunakan aplikasi prediksi calon pegawai yang menerapkan algoritma C4.5.

3. Prediksi Mahasiswa Berpotensi Non Aktif Menggunakan Data Mining dalam Decision Tree dan Algoritma C4.5 (Nur Yanti Lumban Gaol, Sekolah Tinggi Manajemen Ilmu Komputer, STMIK Triguna Dharma Medan, 2019) (Yanti, 2019)

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat diambil kesimpulan. bahwa rule – rule yang dihasilkan menghasilkan kriteria-kriteria yang tepat dalam menganalisa mahasiswa yang berpotensi non aktif. Kriterianya adalah jadwal kuliah, nilai absensi, nilai gagal, pembayaran uang kuliah yang dapat mempengaruhi keputusan pada status mahasiswa aktif dan non aktif.

4. Algoritma C4.5 Berbasis Seleksi Atribut Menentukan Kemungkinan Pengunduran Diri Mahasiswa (Hoiriyah,Fakultas Teknologi Informasi, Universitas Islam Kalimantan Muhammad Arsyad Al Banjari Banjarmasin,2018) (Hoiriyah, Arsyad, 2018)

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat diambil kesimpulan. Dalam penelitian ini dilakukan pengujian model dengan menggunakan Algoritma C4.5 Berbasis Seleksi Atribut dengan menggunakan data Mahasiswa. Model yang dihasilkan

diuji untuk mendapatkan nilai accuracy dari setiap algoritma, dengan menggunakan variable yaitu NIM, Nama Mahasiswa, JK, IPK, Semester, dan Status.

5. Penerapan Algoritma C4.5 Pada Klasifikasi Potensi Siswa Drop Out (Muhamad, Agus Perdana Windarto, Suhada, Program Studi Sistem Informasi STIKOM Tunas Bangsa Pematang Siantar, 2019) (Muhamad, Perdana, Suhada, 2019)

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat diambil kesimpulan. Berdasarkan seluruh hasil tahapan penelitian yang telah dilakukan pada Penerapan Algoritma C4.5 Pada Klasifikasi Siswa Berpotensi Drop Out untuk dapat disimpulkan sebagai berikut. 1. Permasalahan menentukan siswa berpotensi drop out dapat diselesaikan menggunakan teknik data mining, yaitu dengan Algoritma C4.5. menghasilkan 9 (sembilan) rules dan Tingkat akurasi yang dihasilkan oleh metode tersebut adalah 96,15% dan nilai AUC (Area Under the ROC Curve) 0,998. 2. Dengan adanya penerapan data mining algoritma C4.5 diharapkan mampu memberikan solusi dalam menentukan siswa yang berpotensi drop out di MTs. Al-Hikmah Marihat Bandar. Dengan menggunakan variabel yaitu Absensi, Saudara, dan Jarak.

6. Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen (Studi Kasus:Hinet Batam) (Sherly Maisa Putri, Sasa Ani Arnomo, 2020) (Maisa, Ani , 2020)

Kualitas pelayanan merupakan tujuan dari perusahaan dalam memberikan kepuasan kepada konsumen baik perusahaan jasa maupun non jasa. Salah satu perusahaan jasa penyedia layanan akses internet yang berkomitmen untuk meningkatkan kualitas pelayanan terhadap kepuasan konsumen adalah Hinet Batam. Penelitian ini bertujuan untuk menganalisa kualitas pelayanan terhadap kepuasan konsumen Hinet Batam dengan menggunakan algoritma C4.5. Variabel penilaian mencakup empathy, responsiveness, reliability dengan keputusan puas dan tidak puas. Hasil dari pengolahan metode Data Mining dengan menggunakan algoritma C4.5 ternyata mampu mengklasifikasi tingkat kepuasan konsumen terhadap kualitas Pelayanan. Dengan nilai presisi prediksi puas sebesar 61.7% dan nilai presisi prediksi tidakpuas 38,3%. Hasil penelitian ini bisa dijadikan acuan untuk meningkatkan kualitas pelayanansertauntuk mempertahankan loyalitas konsumen dalam menggunakan produk paket data layanan internet dan meningkatkan penjualan.

7. Penerapan Algoritma C4.5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang (Anggita Safitri Febriarini, Erna Zuni Astuti, Program Studi Teknik Informatika Universitas Dian Nuswantoro, 2019) (Safitri, Zuni, Dian, 2019)

Sejumlah aktivitas perkotaan yang ada di Kota Semarang menjadikan tingginya peningkatan yang tentunya membutuhkan dukungan sistem angkutan umum yang andal, cepat, dan efisien. Sebagai bagian dari upaya memecahkan masalah kemacetan, Pemerintah Pusat melalui Departemen Perhubungan membuat Bus Rapid Transit (BRT) yang saat ini diterapkan di berbagai kota di Indonesia. Konsep BRT merupakan sistem angkutan massal yang terintegrasi di setiap koridor, yang bertujuan untuk memenuhi

kebutuhan masyarakat akan transportasi dalam kota. Tujuan penelitian ini adalah untuk mengetahui dan menganalisis kepuasan penumpang BRT Trans Semarang dengan menggunakan teknik data mining dengan metode decision tree tepatnya algoritma C4.5. Dalam penelitian ini variabel pengukuran kepuasan penumpang yang digunakan adalah harga, fasilitas, dan pelayanan. Dari penelitian ini atribut yang digunakan menghasilkan hubungan sebab-akibat dalam mengklasifikasikan penumpang puas dan tidak puas. Penelitian ini diharapkan dapat membantu Pemerintah dalam meningkatkan kepuasan penumpang agar masyarakat beralih menggunakan BRT Trans Semarang dan dapat memberikan pengaruh terjadinya peningkatan jumlah penumpang. Berdasarkan hasil klasifikasi dengan menggunakan algoritma C4.5 menunjukkan adanya peningkatan di tiap pengujian yang telah dilakukan dengan hasil akurasi akhir yang cukup baik sebesar 95% yang menunjukkan bahwa algoritma C4.5 cocok digunakan untuk mengukur tingkat kepuasan penumpang BRT Trans Semarang.

8. **Penerapan Algoritma C4.5 Pada Program klasifikasi Mahasiswa Dropout (Anik Andriani AMIK BSI Jakarta, 2017)** (Andrian, 2017) Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat diambil kesimpulan. 1. Klasifikasi mahasiswa dengan algoritma C4.5 dapat mengklasifikasikan mahasiswa aktif dan dropout. 2. Hasil evaluasi dan validasi dengan confusion matrix menunjukkan tingkat akurasi pada algoritma C4.5 sebesar 97,75%. 3. Hasil evaluasi dan validasi dengan ROC/AUC menunjukkan nilai lebih dari 0,9 sehingga dapat dimasukkan kedalam excellent classification. 4. Penerapan rule dari algoritma C4.5 yang digunakan dalam klasifikasi mahasiswa potensi dropout terhadap data baru diperoleh hasil evaluasi dan validasi dengan confusion matrix menghasilkan tingkat akurasi sebesar 90,00%. Dengan menggunakan variabel Waktu Kuliah, IPK, Kehadiran, Jenis Kelamin, Usia, Asal Daerah, Jurusan SLTA, Orang Tua, Penghasilan Orang Tua, Biaya Studi, Berkerja, Beasiswa.
9. **Penerapan Metode C4.5 Untuk Klasifikasi Mahasiswa Berpotensi Drop Out (Asmaul Husnah Nasrullah, Universitas Ichsan Gorontalo, 2018)** (Husnah, 2018) Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat diambil kesimpulan. 1. Metode C4.5 yang digunakan dalam klasifikasi mahasiswa yang berpotensi drop out menghasilkan tujuh belas (17) rule. 2. 17 rule tersebut dapat dijadikan sebagai pola dalam menentukan mahasiswa yang berpotensi mengalami drop out. Dengan menggunakan variabel Jenis Kelamin, Umur, Agama, Asal Daerah, Kelas, IPS.
10. **Aplikasi Data Mining Menggunakan Algoritma C4.5 Untuk Mengklasifikasi Kelulusan Mahasiswa Pada Universitas Dian Nuswantoro Semarang (Danny Himawan, Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang, 2014)** (Himawan, 2014) Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat diambil kesimpulan pada percobaan dengan 100 dataset dengan 10 sampel uji hasil yang diharapkan tidak mencapai yang diinginkan karena dengan 10 sampel uji 8 sampel tepat dan 2 sampel tidak tepat. Sama halnya dengan 1000 dataset dengan 10 sampel uji hasil yang diharapkan tidak mencapai yang diinginkan

karena dengan 10 sampel uji 8 sampel tepat dan 2 sampel tidak tepat. Sedangkan pada percobaan dengan seluruh dataset / 1469 data dengan 10 sampel uji, semua sampel uji tepat. Sehingga dapat disimpulkan bahwa semakin sedikit data training yang digunakan maka hasil klasifikasinya menunjukkan ketidakakuratan yang tinggi. Sebaliknya jika data training yang digunakan semakin banyak hasilnya akan semakin akurat. Software klasifikasi yang telah disusun dapat digunakan untuk menentukan klasifikasi potensi ketepatan kelulusan pada mahasiswa UDINUS Semarang.

Tabel 2. 1. Tinjauan Studi

NO.	TAHUN	JUDUL	PENYUSUN	KONTRIBUSI / KELEMAHAN
1.	2018	Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada	Muhammad Fauzul Arifin, Devi Fitriyah	Kontribusi terhadap penelitian ini adalah pengambilan metode algoritma C4.5 dilihat dari tingkat akurasi yang dicapai. Kelemahan pada penelitian ini yaitu masih banyaknya aturan-aturan/rule model yg digunakan sebagai acuan dalam perancangan dan pembuatan aplikasi GUI.
2.	2017	Penerapan Algoritma C4.5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE	Fandy Ferdian Harryanto, Seng Hansun	Kontribusi terhadap penelitian ini adalah perhitungan akurasi untuk tiap kluster dan secara keseluruhan. Kelemahan pada penelitian ini kurangnya sampel data yang lebih banyak agar tingkat akurasi aplikasi dapat ditingkatkan.
3	2019	Prediksi Mahasiswa Berpotensi Non Aktif Menggunakan Data Mining dalam Decision Tree dan Algoritma C4.5	Nur Yanti Lumban Gaol	Kontribusi terhadap penelitian ini adalah pengambilan metode Algoritma C4.5 dan cara perhitungan C4.5 dalam prediksi pengambilan variabel. Kelemahan pada penelitian ini yaitu mendeteksi status mahasiswa menggunakan Aplikasi.
4	2018	Algoritma C4.5 Berbasis Seleksi Atribut	Hoiriyah	Kontribusi terhadap penelitian ini adalah mengenai hasil rule yang didapat dari penelitian tersebut.

		Menentukan Kemungkinan Pengunduran Diri Mahasiswa		Kelemahan pada penelitian ini yaitu mendeteksi status mahasiswa menggunakan Aplikasi.
5	2019	Penerapan Algoritma C4.5 Pada Klasifikasi Potensi Siswa Drop Out	Muhamad, Agus Perdana Windarto, Suhada,	Kontribusi terhadap penelitian ini adalah pengambilan metode Algoritma C4.5 dan cara perhitungan C4.5. Kelemahan pada penelitian ini yaitu belum adanya perhitungan nilai akurasi.
6	2020	Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen(Studi Kasus:Hinet Batam)	Sherly Maisa Putri, Sasa Ani Arnomo	Kontribusi terhadap penelitian ini adalah Hasil dari pengolahan metode Data Mining dengan menggunakan algoritma C4.5 ternyata mampu dalam melakukan klasifikasi data karena karakteristik data yang diklasifikasi dapat diperoleh dengan jelas,baik dalam struktur pohon keputusan maupun aturan if-then. Kelemahan pada penelitian ini yaitu dalam menggunakan produk paket data layanan internet dalam meningkatkan penjualan.
7	2019	Penerapan Algoritma C4.5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang	Anggita Safitri Febriarini, Erna Zuni Astuti	Kontribusi terhadap penelitian ini adalah pengujian yang dilakukan, untuk mengetahui nilai akurasi,precision dan recall Kelemahan pada penelitian ini yaitu Penelitian ini bisa dikembangkan dengan menggabungkan atau membandingkan dengan algoritma prediksi lain untuk memperoleh hasil prediksi yang lebih baik.
8	2017	Penerapan Algoritma C4.5 Pada Program	Anik Andriani	Kontribusi terhadap penelitian adalah pengambilan uji hasil menggunakan confusion matrix

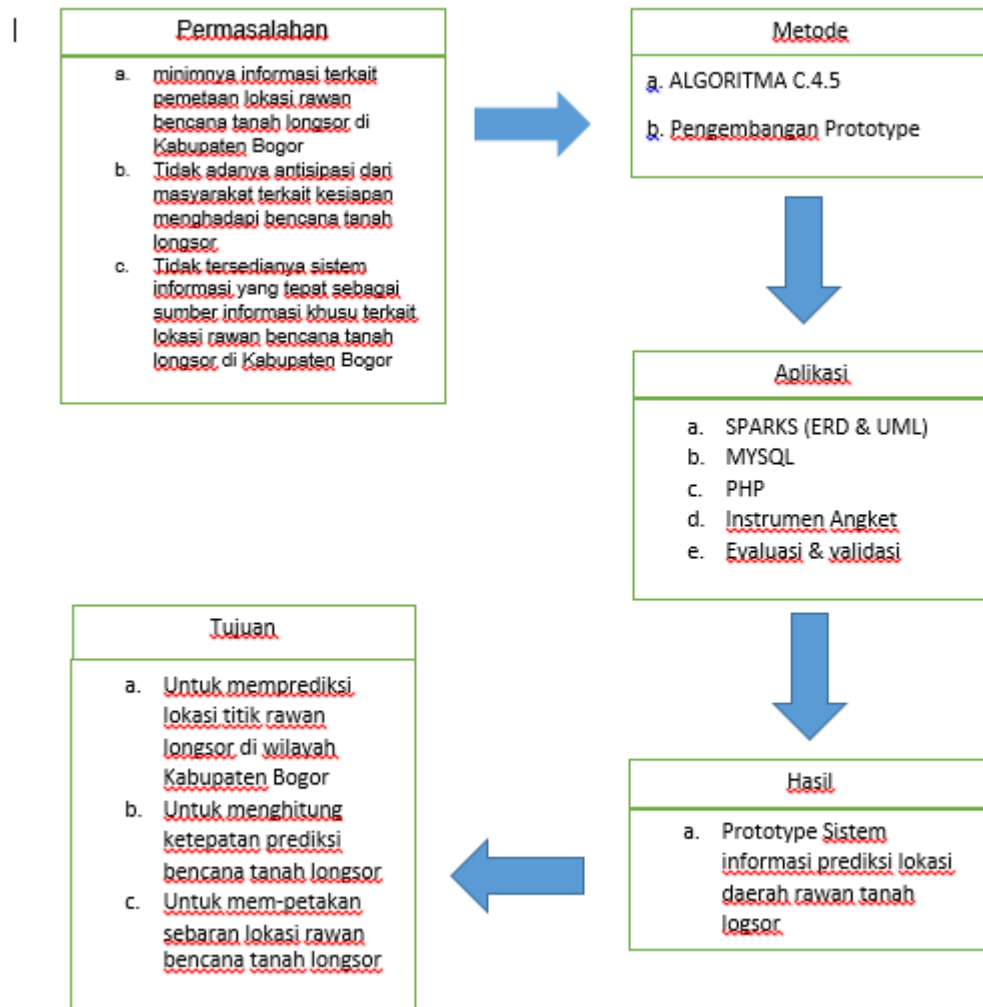
		klasifikasi Mahasiswa Dropout		Kelemahan pada penelitian ini yaitu tidak terdapat cara perhitungan Algoritma C4.5.
9	2018	Penerapan Metode C4.5 Untuk Klasifikasi Mahasiswa Berpotensi Drop Out	Asmaul Husnah Nasrullah	Kontribusi terhadap penelitian ini adalah mengenai hasil rule yang didapat dari penelitian tersebut. Kelemahan pada penelitian ini yaitu penentuan pola yang ditemukan tidak dibuatkan dalam aplikasi .
10	2014	Aplikasi Data Mining Menggunakan Algoritma C4.5 Untuk Mengklasifikasi Kelulusan Mahasiswa Pada Universitas Dian Nuswantoro Semarang	Danny Himawan	Kontribusi terhadap penelitian ini adalah pengambilan evaluasi usability menggunakan blackbox. Kelemahan pada penelitian ini yaitu tidak terdapat cara perhitungan Algoritma C4.5.

Sesuai tabel yang diatas yang menjadi acuan pada penelitian ini adalah Jurnal atau sumber dari nomor 2 dan 3. Dalam penelitian yang menjadi acuan metodenya adalah nomor 5.

Metode dan hasil dari 10 jurnal tersebut menggunakan metode algoritma C4.5 yang dapat digunakan untuk prediksi. Sedangkan untuk persamaan dengan penelitian ini yaitu menggunakan metode algoritma C4.5 dapat digunakan untuk mengetahui potensi lokasi tanah longsor.

D. Kerangka Pemikiran

Kerangka pemikiran ini dibuat mewakili konsep pemecahan masalah penelitian yang meliputi masalah yang akan di teliti, metode penelitian, dan hasil dari penelitian, berikut dijelaskan berdasarkan gambar 2.5



Gambar 2. 10 Kerangka Pemikiran

Keterangan :

1. Masalah dalam penelitian ini yaitu minimnya informasi terkait sebaran lokasi rawan bencana tanah longsor sehingga mengurangi tingkat antisipasi dan penanggulangan bencana terhadap masyarakat dan belum adanya system informasi khusus sebagai wadah sebaran lokasi rawan bencana tanah longsor.
2. Tujuan penelitian ini yaitu, untuk memprediksi lokasi rawan tanah longsor dan menghitung ketepatan prediksi serta mempetakan sebaran wilayah lokasi rawan bencana tanah. Metode penelitian pengembangan ini menggunakan metode algoritma

C.4.5 sebagai dasar analisis data penelitian dan pengembangan yang digunakan dalam pembangunan system nantinya menggunakan metode *prototype*.

3. Aplikasi yang digunakan untuk pembangunan sistem prototype dalam penelitian pengembangan ini menggunakan SPARK sebagai aplikasi penunjang untuk membuat diagram UML dan ERD, desain database menggunakan MySQL, pemrograman menggunakan bahasapemrograman PHP. Sebaran angket sebagai proses pengujian system prototype dan implementasi sebagai dasar evaluasi dan validasi system prototype yang akan dibuat nantinya.
4. Hasil yang diharapkan dari penelitian ini yaitu menghasilkan system informasi prototype khusus sebagai informasi terkait lokasi rawan tanah longsor, yang nantinya dapat digunakan oleh stakeholder terkait sebagai tindak lanjut untuk pengantisipasi terjadinya bencana tanah longsor di Kabupaten Bogor,

E. Hipotesis

Hipotesis penelitian adalah jawaban sementara terhadap pertanyaan-pertanyaan penelitian. Hipotesis dapat dijelaskan dari berbagai sudut pandang, misalnya secara etimologis, teknis, statistik, dan lain sebagainya. Hipotesis dalam penelitian ini adalah Metode Algoritma C4.5 diduga dapat memprediksi kualitas air sungai.