

## PENERAPAN DATA MINING DALAM MEMPREDIKSI PEMENANG KLUB SEPAK BOLA PADA AJANG LIGA CHAMPION DENGAN ALGORITMA C.45

Yonata Laia\*, Charles Tandian, Andi Saputra

Program Studi Sistem Informasi, Fakultas Teknologi dan Ilmu Komputer, Universitas Prima Indonesia

E-mail : \*yonata@unprimdn.ac.id

**ABSTRAK** - Liga Champion adalah sebuah kompetisi sepak bola antar klub tahunan yang diikuti oleh klub divisi tertinggi Eropa. Kompetisi ini merupakan salah satu turnamen paling bergengsi di dunia dan paling diminati oleh seluruh penggemar sepak bola dunia. Dengan diikuti oleh klub divisi tertinggi di Eropa, kompetisi ini sangat susah untuk ditebak klub mana yang akan unjuk gigi di akhir pertandingan kompetisi tersebut. Hanya bermodalkan bakat, fisik, dan *skill* belum tentu bisa memenangkan kompetisi bergengsi tersebut. Penelitian akan membahas bagaimana merancang suatu *data mining* untuk memprediksi klub sepak bola manakah yang akan memenangkan kompetisi Liga Champion ini. Teknik *data mining* yang digunakan adalah algoritma C.45 yang mana metode ini membuat sebuah pohon keputusan yang sangat terkenal dalam mengklasifikasi dan memprediksi. Dari hasil penelitian ini, dapat diketahui bahwa dengan menggunakan *data mining* teknik algoritma C.45 dapat membantu penggemar sepak bola dalam memprediksi klub manakah yang akan menjadi pemenang kompetisi Liga Champion.

Kata kunci: prediksi, *data mining*, sepak bola, algoritma C.45

### 1. PENDAHULUAN

Sepakbola adalah olahraga berskala internasional, permainan tersebut membangkitkan luapan keinginan dan emosi yang tidak sama dengan olahraga lainnya. Sepakbola merupakan sesuatu yang umum di antara orang-orang dengan latar belakang dan keturunan yang berbeda-beda, sebuah jembatan yang menghubungkan jejang ekonomi, politik, kebudayaan, dan agama. Dikenal sebagai bola kaki hampir di seluruh dunia, sepak bola merupakan olahraga nasional hampir di seluruh negara di Asia, Afrika, Eropa, dan Amerika Selatan. Olahraga ini merupakan satu-satunya jenis permainan bola kaki yang dimainkan di olimpiade.

Dalam perkembangannya, sepak bola tidak dianggap sebagai sebuah cabang olahraga saja, tetapi sepak bola telah menjadi salah satu dimensi kuat dalam kehidupan manusia yang menggemari sepak bola. Pada era sepak bola modern, sepak bola sudah dianggap sebagai sebuah hiburan bahkan sepak bola modern sudah dijadikan sebagai sebuah industri atau sebuah bisnis yang dianggap dapat mendatangkan keuntungan yang besar bagi pemilik klubnya.

Prediksi bola, sering sekali menjadi bahan perbincangan hangat dari masyarakat pecinta bola. Setiap pecinta bola pasti ingin sekali menyaksikan partai-partai bola yang enak ditonton dan seru. Apabila ketika tim kesayangan kita yang akan bermain, pastinya hanya ada satu hasil yang diinginkan, yaitu menang<sup>[1][2][3]</sup>.

Liga Champions UEFA adalah sebuah kompetisi sepak bola antarklub tahunan yang diselenggarakan oleh Uni Sepak Bola Eropa/*Union of European Football Associations* (UEFA) dan diikuti oleh klub divisi tertinggi Eropa. Kompetisi

ini merupakan salah satu turnamen paling bergengsi di dunia dan kompetisi antarklub paling bergengsi di sepak bola Eropa, yang hanya diikuti oleh juara liga nasional (dan juga juara kedua untuk beberapa negara) dari setiap asosiasi nasional anggota UEFA. Final Liga Champions UEFA adalah acara yang paling banyak ditonton di seluruh dunia setiap tahunnya. Final musim 2012–13 merupakan yang paling banyak ditonton, dengan jumlah mencapai 360 juta penonton televisi. Sejak tahun 2015, Final Liga Champions UEFA digelar pada hari Sabtu minggu pertama bulan Juni pada tahun ganjil dan Sabtu minggu terakhir bulan Mei pada tahun genap<sup>[4]</sup>.

Hasil pertandingan sepakbola merupakan suatu hal yang tidak bisa diperkirakan dengan pasti, tetapi dari data-data pertandingan dapat diketahui adanya suatu pola yang dapat digunakan untuk memperkirakan hasil dari suatu pertandingan. Kebutuhan akan peta kekuatan calon lawan dari tim yang akan bertanding, dengan tujuan untuk meraih kemenangan sangat beralasan. Dengan mengetahui peta kekuatan lawan, sebuah tim khususnya seorang pelatih dapat mempersiapkan dan mengantisipasi kekuatan tim calon lawannya serta dapat memfikirkan sebuah strategi dan instruksi-instruksi khusus kepada para pemainnya. Melalui sistem ini pengguna dapat melihat prediksi hasil pertandingan sepak bola dan kekuatan dari masing-masing tim yang akan bertanding.

Karena banyaknya komentar-komentar yang dilontarkan, banyak yang mengatakan prediksi bola pasti. Tidak ada satu orangpun yang dapat memprediksi bola pasti menang, pasti kalah atau pun pasti seri atau mengetahui apa yang terjadi sebelum pluit berbunyi tanda berakhirnya pertandingan.

Memang banyak faktor yang mempengaruhi dalam menentukan prediksi sepakbola, penelitian ini bertujuan untuk menyelesaikan masalah diatas, yaitu mencoba menggunakan algoritma C.45 untuk menggali informasi dari data pertandingan sebelumnya untuk melakukan prediksi pertandingan sepakbola.

Oleh karena itu, maka dalam penelitian ini penulis membuat *data mining* dengan metode algoritma C.45 dengan harapan seluruh penggemar sepak bola dapat memprediksi klub pemenang Liga Champion dengan cara yang efektif dan efisien<sup>[5][6][7]</sup>.

## 2. ISI PENELITIAN

### 2.1 Metodologi

*Data Mining* merupakan proses penggalian dan pertambangan pengetahuan dari sejumlah data yang besar, *database*, atau *repository database* lainnya. Tujuan utama dari penambangan data ini untuk menemukan pengetahuan baru yang tersembunyi dari database tersebut.

Pekerjaan yang berkaitan dengan data mining dapat dibagi menjadi empat kelompok, yaitu model prediksi (*prediction modelling*), analisis kelompok (*cluster analysis*), analisis asosiasi (*association analysis*), dan deteksi anomali (*anomaly detection*).

#### 1. Model Prediksi

Model prediksi berkaitan dengan pembuatan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat. Ada dua jenis model prediksi, yaitu klasifikasi dan regresi. Klasifikasi digunakan untuk variabel target diskret, sedangkan regresi untuk variabel target kontinu.

#### 2. Analisis Kelompok

Analisis kelompok melakukan pengelompokan data - data kedalam sejumlah kelompok (*cluster*) berdasarkan kesamaan karakteristik masing - masing data pada kelompok - kelompok yang ada. Data - data yang masuk dalam batas kesamaan dengan kelompoknya akan bergabung dalam kelompok tersebut, dan akan terpisah dalam kelompok yang berbeda jika keluar dari batas kesamaan dengan kelompok tersebut.

#### 3. Analisis Asosiasi

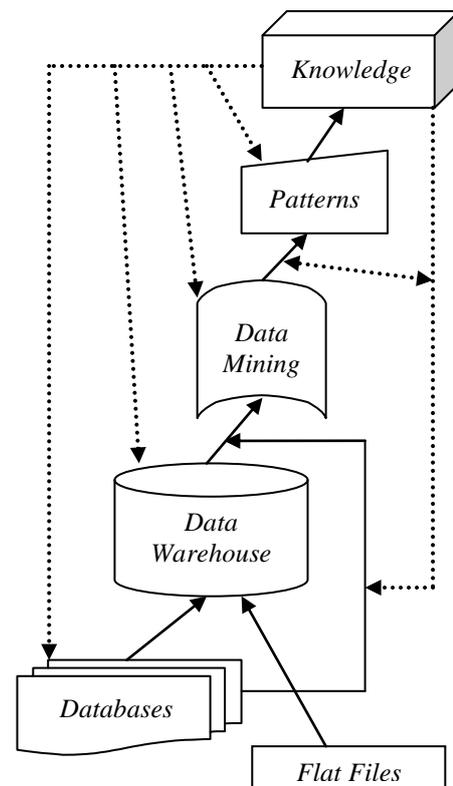
Analisis asosiasi digunakan untuk menemukan pola yang menggambarkan kekuatan hubungan fitur dalam data. Pola yang ditemukan biasanya merepresentasikan bentuk aturan implikasi atau subset fitur. Tujuannya adalah untuk menemukan pola yang menarik dengan cara yang efisien.

#### 4. Deteksi Anomali

Deteksi anomali berkaitan dengan pengamatan sebuah data dari sejumlah data yang secara signifikan mempunyai karakteristik yang berbeda dari sisa data yang lain. Data - data yang

karakteristiknya menyimpang (berbeda) dari data lain disebut *outliner*. Algoritma deteksi anomali yang baik harus mempunyai laju deteksi yang tinggi dan laju *error* yang rendah. Deteksi anomali dapat diterapkan pada sistem jaringan untuk mengetahui pola data yang memasuki jaringan sehingga penyusupan bisa ditemukan jika pola kerja data yang datang berbeda. Perilaku kondisi cuaca yang mengalami anomali juga dapat dideteksi dengan algoritma ini<sup>[8][9]</sup>.

Karena *Data Mining* adalah suatu rangkaian proses yang dapat dibagi menjadi beberapa tahap. Tahap - tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantara *knowledge base*. Tahap - tahap ini diilustrasikan pada gambar berikut :



Gambar 1. Tahapan *Data Mining*

Di dalam kehidupan manusia sehari - hari, manusia selalu dihadapkan oleh berbagai macam masalah dari berbagai macam bidang. Masalah - masalah yang dihadapi oleh manusia memiliki tingkat kesulitan dan kompleksitas yang sangat bervariasi, mulai dari masalah yang teramat sederhana dengan sedikit faktor - faktor yang terkait, sampai dengan masalah yang sangat rumit dengan banyak sekali faktor - faktor yang terkait dan perlu untuk diperhitungkan. Untuk menghadapi masalah - masalah ini, manusia mulai mengembangkan sebuah sistem yang dapat membantu manusia agar dapat dengan mudah mampu untuk menyelesaikan masalah - masalah tersebut. Adapun pohon keputusan ini adalah

sebuah jawaban akan sebuah sistem yang manusia kembangkan untuk membantu mencari dan membuat keputusan untuk masalah – masalah tersebut dan dengan memperhitungkan berbagai macam faktor yang ada di dalam lingkup masalah tersebut. Dengan pohon keputusan, manusia dapat dengan mudah mengidentifikasi dan melihat hubungan antara faktor - faktor yang mempengaruhi suatu masalah dan dapat mencari penyelesaian terbaik dengan memperhitungkan faktor - faktor tersebut. Pohon keputusan ini juga dapat menganalisa nilai resiko dan nilai suatu informasi yang terdapat dalam suatu alternatif pemecahan masalah. Peranan pohon keputusan sebagai alat bantu dalam mengambil keputusan (*decision support tool*) telah dikembangkan oleh manusia sejak perkembangan teori pohon yang 54 dilandaskan pada teori graf. Kegunaan pohon keputusan yang sangat banyak ini membuatnya telah dimanfaatkan oleh manusia dalam berbagai macam sistem pengambilan keputusan.

Pohon dalam analisis pemecahan masalah pengambilan keputusan adalah pemetaan mengenai alternatif – alternatif pemecahan masalah yang dapat diambil dari masalah tersebut. Pohon tersebut juga memperlihatkan faktor - faktor kemungkinan probabilitas yang akan mempengaruhi alternatif - alternatif keputusan tersebut, disertai dengan estimasi hasil akhir yang akan didapat bila kita mengambil alternatif keputusan tersebut<sup>[10]</sup>.

Arsitektur pohon keputusan dibuat sedemikian rupa agar menyerupai pohon asli, dimana terdapat beberapa bagian yaitu:

1. *Root Node*: *Node* ini terletak pada bagian paling atas dari pohon keputusan.
2. *Internal Node*: *Node* ini merupakan percabangan dimana membutuhkan satu input dan mengeluarkan maksimal dua *output*.
3. *Leaf Node*: *Node* ini merupakan *node* yang terletak pada ujung pohon. *Node* ini hanya memiliki satu input dan tidak memiliki output.

Kelebihan dari metode pohon keputusan adalah:

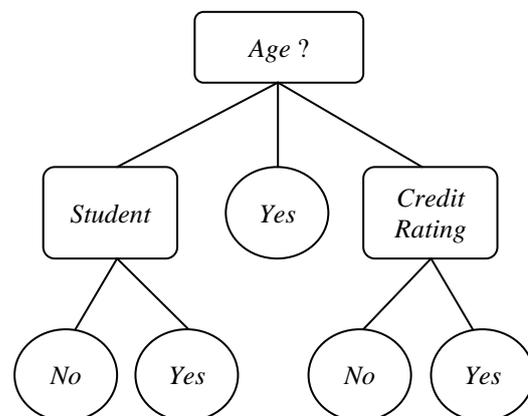
1. Daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat diubah menjadi lebih simpel dan spesifik.
2. Eliminasi perhitungan - perhitungan yang tidak diperlukan, karena ketika menggunakan metode pohon keputusan maka sampel diuji hanya berdasarkan kriteria atau kelas tertentu.
3. Fleksibel untuk memilih fitur dari internal node yang berbeda, fitur yang terpilih akan membedakan suatu kriteria dibandingkan kriteria yang lain dalam node yang sama. Kefleksibelan metode pohon keputusan ini meningkatkan kualitas keputusan yang dihasilkan jika dibandingkan ketika menggunakan metode penghitungan satu tahap yang lebih konvensional.

4. Dalam analisis multivariat, dengan kriteria dan kelas yang jumlahnya sangat banyak, seorang penguji biasanya perlu untuk mengestimasi baik itu distribusi dimensi tinggi ataupun parameter tertentu dari distribusi kelas tersebut. Metode pohon keputusan dapat menghindari munculnya permasalahan ini dengan menggunakan kriteria yang jumlahnya lebih sedikit pada setiap node internal tanpa banyak mengurangi kualitas keputusan yang dihasilkan.

Kekurangan dari metode pohon keputusan:

1. Terjadi *overlap* terutama ketika kelas – kelas dan kriteria yang digunakan jumlahnya sangat banyak. Hal tersebut juga dapat menyebabkan meningkatnya waktu pengambilan keputusan dan jumlah memori yang diperlukan.
2. Pengakumulasian jumlah *error* dari setiap tingkat dalam sebuah pohon keputusan yang besar.
3. Kesulitan dalam mendesain pohon keputusan yang optimal.
4. Hasil kualitas keputusan yang didapatkan dari metode pohon keputusan sangat tergantung pada bagaimana pohon tersebut didesain.

Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Contoh dari pohon keputusan dapat dilihat di Gambar 2 berikut ini.



Gambar 2. Model Pohon Keputusan

Disini setiap percabangan menyatakan kondisi yang harus dipenuhi dan tiap ujung pohon menyatakan kelas data. Contoh di Gambar 2. adalah identifikasi pembeli komputer, dari pohon keputusan tersebut diketahui bahwa salah satu kelompok yang potensial membeli komputer adalah orang yang berusia di bawah 30 tahun dan juga pelajar. Setelah sebuah pohon keputusan dibangun maka dapat digunakan untuk mengklasifikasikan *record* yang belum ada kelasnya. Dimulai dari *node root*, menggunakan tes terhadap atribut dari *record* yang belum ada kelasnya tersebut lalu mengikuti cabang yang sesuai dengan hasil dari tes tersebut, yang akan membawa kepada *internal node* (*node*

yang memiliki satu cabang masuk dan dua atau lebih cabang yang keluar), dengan cara harus melakukan tes lagi terhadap atribut atau *node* daun. *Record* yang kelasnya tidak diketahui kemudian diberikan kelas yang sesuai dengan kelas yang ada pada *node* daun. Pada pohon keputusan setiap simpul daun menandai label kelas. Proses dalam pohon keputusan yaitu mengubah bentuk data (tabel) menjadi model pohon (tree) kemudian mengubah model pohon tersebut menjadi aturan (*rule*).

*Knowledge Discovery In Databases* ( KDD ) adalah proses non trivial untuk mencari dan mengidentifikasi pola (*pattern*) dalam data. *Knowledge Discovery In Database* (KDD) merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola, atau hubungan dalam data berukuran besar serta hubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi, dan visualisasi dari pola - pola sejumlah kumpulan data.

Secara umum tahapan – tahapan proses *Knowledge Discovery in Databases* terdiri dari:

1. *Data Cleaning*  
Proses menghilangkan *noise* dari data yang tidak konsisten.
2. *Data Integration*  
Penggabungan data dari berbagai *database* ke dalam satu *database* baru.
3. *Data Selection*  
Proses pemilihan data yang relevan yang didapat dari *database*.
4. *Data Transformation*  
Data diubah ke dalam format yang sesuai untuk diproses dalam *Data Mining*.
5. *Data Mining*  
Suatu metode yang diterapkan untuk menemukan pengetahuan berharga yang tersembunyi dari data.
6. *Pattern Evaluation*  
Mengidentifikasi pola - pola menarik untuk dipresentasikan ke dalam *knowledge based*.
7. *Knowledge Presentation*  
Visualisasi dan penyajian pengetahuan mengenai teknik yang digunakan untuk memperoleh pengetahuan yang diperoleh oleh *user*.

Algoritma C.45 yaitu sebuah algoritma yang digunakan untuk membangun *decision tree* (pengambilan keputusan). Algoritma C.45 adalah salah satu algoritma induksi pohon keputusan, yaitu ID3 (*Iterative Dichotomiser 3*). ID3 dikembangkan oleh J. Ross Quinlan. Dalam prosedur algoritma ID3, input berupa sampel *training*, label training, dan atribut. Algoritma C.45 merupakan pengembangan dari ID3.

Pohon dibangun dengan cara membagi data secara rekursif hingga tiap bagian terdiri dari data

yang berasal dari kelas yang sama. Bentuk pemecahan (*split*) yang digunakan untuk membagi data tergantung dari jenis atribut yang digunakan dalam *split*. Algoritma C4.5 dapat menangani data numerik (kontinyu) dan diskret. *Split* untuk atribut numerik yaitu mengurutkan contoh berdasarkan atribut kontinyu A, kemudian membentuk minimum permulaan (*threshold*) M dari contoh - contoh yang ada dari kelas mayoritas pada setiap partisi yang bersebelahan, lalu menggabungkan partisipartisi yang bersebelahan tersebut dengan kelas mayoritas yang sama. *Split* untuk atribut diskret A mempunyai bentuk *value* (A) X dimana  $X \subset \text{domain}(A)$ . Jika suatu set data mempunyai beberapa pengamatan dengan *missing value* yaitu *record* dengan beberapa nilai variabel tidak ada, Jika jumlah pengamatan terbatas maka atribut dengan *missing value* dapat diganti dengan nilai rata - rata dari variabel yang bersangkutan. Dalam melakukan pemisahan obyek (*split*) dilakukan tes terhadap atribut dengan mengukur tingkat ketidakmurnian pada sebuah simpul (*node*)<sup>[11][12]</sup>.

Secara singkat logika algoritma C4.5 yang digunakan adalah sebagai berikut:

1. Pilih atribut sebagai akar
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk masing - masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C.45, yaitu :

1. Menyiapkan data *training*. Data *training* biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas – kelas tertentu .
2. Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai *gain* dari masing – masing atribut, nilai *gain* yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai *gain* dari atribut, hitung dahulu nilai *entropy*. Untuk menghitung nilai *entropy* digunakan rumus

$$Entropy(S) = - \sum_{i=1}^n - p_i \cdot \log_2 p_i \quad (1)$$

Keterangan :

S = himpunan kasus

n = jumlah partisi S

Pi = proporsi Si terhadap S

3. Menghitung nilai *gain* menggunakan rumus:  
 $Gain(A) =$

$$Entropy(S) - \sum_{i=1}^n |S_i|/|S| \times Entropy(S_i) \quad (2)$$

Keterangan :

S = himpunan kasus

- A = atribut  
 n = jumlah partisi atribut A  
 $|S_i|$  = jumlah kasus pada partisi ke i  
 $|S|$  = jumlah kasus dalam S

4. Ulangi langkah ke-2 hingga semua *record* terpartisi.
5. Proses partisi pohon keputusan akan berhenti saat :
  - a) Semua *record* dalam simpul N mendapat kelas yang sama.
  - b) Tidak ada atribut di dalam *record* yang dipartisi lagi.
  - c) Tidak ada *record* di dalam cabang yang kosong.

## 2.2 Hasil

Penelitian ini akan mengidentifikasi faktor – faktor yang umumnya menjadi syarat kemenangan suatu klub pada ajang suatu kompetisi. Yang menjadi *input* dalam penelitian ini adalah nama klub yang mengikuti kompetisi tersebut sedangkan outputnya berupa nama klub yang merupakan kandidat terkuat yang akan memenangkan kompetisi Liga Champion tersebut. *Output* atau target variabel memiliki dua kategori yaitu Ya dan Tidak, maka digunakan teknik klasifikasi yang termasuk dalam pengelompokan *data mining*, karena dalam klasifikasi terdapat target variabel kategori.

Semua data sepak bola akan dicari hubungannya dengan data tim bola, hanya beberapa atribut yang kira - kira berguna dan datanya tidak terlalu acak. Karena data yang acak akan membuat proses *mining* memakan waktu yang lama dan tingkat hubungannya pun rendah. Data sepak bola yang akan dicari hubungannya meliputi proses Kualitas Pemain, Umur, Kualitas Pelatih, Finansial, dan Prestasi Klub. Adapun variabel yang akan diproses *mining* meliputi :

### 1. Kualitas Pemain

Hal - hal yang harus dinilai untuk menjadi pemain yang berkualitas adalah sebagai berikut :

- Kemampuan menggiring dan mengontrol bola
- Baik Dalam Bekerja Sama
- Moral dan Kedisiplinan
- Mengikuti Sesi Latihan yang Rutin
- Mental dan Stamina yang Baik
- Komunikasi yang Baik Antar Pemain Dengan Manajemen Klub

### 2. Umur

Hal - hal yang harus dinilai dalam kategori umur adalah sebagai berikut :

- Pengalaman Bermain Sepak Bola
- Usia Menginjak Kepala 3
- Umur yang produktif
- Usia di Bawah 20 Tahun
- Rata – Rata Usia Keseluruhan Pemain Dalam Tim

### 3. Kualitas Pelatih

Hal - hal yang harus dinilai untuk menjadi pelatih yang berkualitas adalah sebagai berikut :

- Pandai Dalam Membelanjakan Pemain
- Konsisten Dalam Permainan
- Strategi yang anti monoton
- Usia yang masih produktif
- Kedisiplinan Dalam Melatih
- Bertanggung Jawab Apabila Tim Sering Mengalami Kekalahan
- Ketegasan
- Hubungan yang Baik Dengan Manajemen Klub

### 4. Finansial

Hal - hal yang harus dinilai dalam kategori finansial adalah sebagai berikut :

- Kekayaan Klub
- Gaji Pemain yang Melimpah
- Staff dan Petinggi yang Bagus
- Mempunyai Banyak Pemain Dengan Transfer Value yang Tinggi
- Tempat Latihan yang Nyaman
- Stadium Kandang yang Berkelas Tinggi
- Fasilitas Pemain yang Memukau

### 5. Banyaknya Prestasi

Hal - hal yang harus dinilai dalam kategori Prestasi Klub adalah sebagai berikut :

- *Trophy* yang Melimpah
- Banyak Melahirkan pemain *Legend*
- Prestasi Bagus yang Turun Temurun
- Selalu Berada di Urutan Teratas di Setiap Liga
- Konsisten Mengikuti Ajang Kompetisi Liga Champion Setiap Tahunnya
- Banyak Penggemar

Data awal *numeric* maupun *nonnumeric* akan dibagi perkelas untuk memudahkan analisa berikutnya. Setelah semua data yang akan dimasukkan dibagi perkelas, maka dilakukan proses klasifikasi dengan membuat pohon keputusan sebagai *output*.

Proses pengambilan keputusan untuk mengidentifikasi faktor - faktor Kemenangan Suatu Klub adalah sebagai berikut :

1. Kualitas Pemain
2. Umur
3. Kualitas Pelatih
4. Finansial
5. Banyaknya Prestasi

Variabel yang akan menjadi keputusan adalah IYA dan TIDAK.

Pada Tabel 1. di bawah ini implementasi perhitungan kasus Algoritma C.45 untuk menentukan kemenangan klub Sepak bola berdasarkan 4 klasifikasi, yaitu kualitas Pelatih, umur, kualitas pemain, dan finansial. Klasifikasi ini dijadikan atribut awal dalam implementasi algoritma C.45 dan atribut tujuannya adalah

peluang dengan klasifikasi dari atribut peluang adalah Iya dan Tidak.

Tabel 1. Data Untuk Klasifikasi Sepak Bola

No	Nama Klub	Kualitas Pemain	Umur
1	Man Utd	Bagus	Rata - Rata
2	FC Basel	Menengah	Muda
3	PSG	Bagus	Rata - Rata
4	Bayen	Bagus	Rata - Rata
5	Roma	Menengah	Muda
6	Chelsea	Bagus	Rata - Rata
7	Barcelona	Bagus	Rata - Rata
8	Juventus	Bagus	Rata - Rata
9	Liverpool	Bagus	Muda
10	Man. City	Bagus	Muda
11	Napoli	Bagus	Tua
12	Besiktas	Menengah	Tua
13	FC Porto	Menengah	Muda
14	Tottenham	Bagus	Rata - Rata
15	Real Madrid	Bagus	Tua
16	Sevilla	Menengah	Muda

Tabel 2. Pembentukan Pohon Keputusan

Kualitas Pelatih	Finansial	Peluang Tinggi
Baru Melatih	Tinggi	Iya
Baru Melatih	Rendah	Tidak
Baru Melatih	Tinggi	Iya
Berprestasi	Tinggi	Iya
Baru Melatih	Rendah	Tidak
Baru Melatih	Tinggi	Iya
Baru Melatih	Tinggi	Iya
Berprestasi	Tinggi	Iya
Berprestasi	Tinggi	Iya
Berprestasi	Tinggi	Iya
Baru Melatih	Rendah	Tidak
Baru Melatih	Rendah	Tidak
Baru Melatih	Rendah	Tidak
Berprestasi	Tinggi	Iya
Baru Melatih	Tinggi	Tidak
Berprestasi	Rendah	Tidak

Tabel 2. berikut ini adalah penjelasan lebih terperinci mengenai tiap - tiap langkah dalam pembentukan pohon keputusan dengan menggunakan algoritma C.45. Entropy adalah ukuran dari teori informasi yang dapat mengetahui karakteristik. Salah satu proses kalkulasi dari entropy adalah proses kalkulasi nilai entropy kelancaran.

Tabel 3. Tabel Klasifikasi Kriteria Node 1

Node	Jumlah Kasus	Iya	Tidak
1	Total	16	9
	Kualitas Pemain :		
	Bagus	11	2
	Menengah	5	5

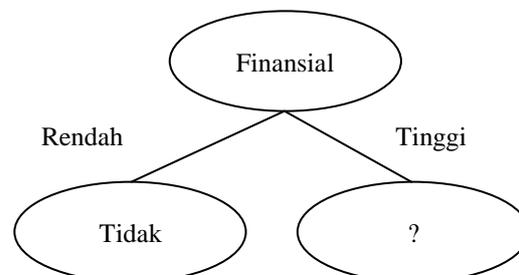
	Umur :			
	Muda	6	2	4
	Rata - Rata	7	7	0
	Tua	3	0	3
	Kualitas Pelatih :			
	Baru Melatih	10	4	6
	Berprestasi	6	5	1
	Finansial :			
	Tinggi	10	9	1
	Rendah	6	0	6

Pohon keputusan dibuat setelah menghitung entropy total, entropy masing - masing atribut dan menghitung gain dan menentukan gain tertinggi. Kunci pencarian Entrophy :

1. Jika diantara kolom "Ya" atau "Tidak" ada yang bernilai 0 maka entropy-nya dipastikan juga bernilai 0.
2. Jika kolom "Ya" dan "Tidak" mempunyai nilai yang sama maka entropy-nya dipastikan juga bernilai 1.

Tabel 4. Hasil Perhitungan Node 1

Node	Jumlah Kasus	Entropy	Gain
1	Total	16	0,988699
	Kualitas Pemain :		
	Bagus	11	0,684038
	Menengah	5	0
	Umur :		
	Muda	6	0,918296
	Rata - Rata	7	0
	Tua	3	0
	Kualitas Pelatih :		
	Baru Latih	10	0,970951
	Berprestasi	6	0,650022
	Finansial :		
	Tinggi	10	0,468996
	Rendah	6	0



Gambar 3. Pohon Keputusan Hasil Perhitungan Node 1

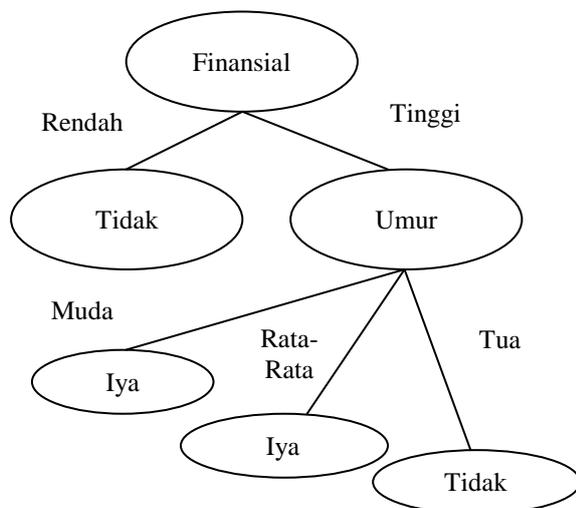
Selanjutnya dihitung nilai *Entropy* dan *Gain* untuk masing - masing atribut untuk dapat menjadi *node* akar dari atribut Tinggi, sama dengan cara yang di atas dengan menghitung nilai *entropy* dari atribut yang tersisa yaitu kualitas pemain, umur, dan kualitas pelatih.

Tabel 5. Tabel Klasifikasi Kriteria *Node* 1.1

Node	Jumlah Kasus	Iya	Tidak
1.1	Total	10	9
	Kualitas Pemain :		
	Bagus	10	9
	Menengah	0	0
	Umur :		
	Muda	2	2
	Rata - Rata	7	7
	Tua	1	0
	Kualitas Pelatih :		
	Baru Melatih	5	4
	Berprestasi	5	5

Tabel 6. Hasil Perhitungan *Node* 1.1

Node	Jumlah Kasus	Entropy	Gain
1.1	Total	10	0,468996
	Kualitas Pemain :		0
	Bagus	10	0,468996
	Menengah	0	0
	Umur :		0,468996
	Muda	2	0
	Rata - Rata	7	0
	Tua	1	0
	Kualitas Pelatih :		0,108032
	Baru Latih	5	0,721928
	Berprestasi	5	0



Gambar 4. Pohon Keputusan *Node* 1.1

Dari hasil pada Tabel 5. dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah Umur, yaitu sebesar 0,468996. Dengan demikian Umur dapat menjadi *node* cabang dari nilai atribut Tinggi. Ada tiga nilai atribut dari Umur, yaitu Muda, Rata - Rata, dan Tua. Dari tiga atribut tersebut, nilai atribut Muda sudah mengklasifikasi kasus menjadi 1, yaitu keputusannya Iya, nilai atribut Rata - Rata juga sudah mengklasifikasi kasus menjadi 1, yaitu keputusannya Iya serta nilai atribut Tua juga sudah mengklasifikasi kasus menjadi 1, yaitu keputusannya Tidak, sehingga tidak perlu dilakukan perhitungan lebih lanjut. Dengan demikian diperoleh pohon keputusan tampak seperti pada Gambar 4.

### 3. KESIMPULAN

Dari uraian pada bab - bab yang sudah dibahas sebelumnya dapat ditarik kesimpulan :

1. Hasil dari pengolahan metode *Data Mining* dengan algoritma C.45 untuk mengidentifikasi kriteria - kriteria kemenangan klub dalam kompetisi Liga Champion adalah kondisi manajemen finansial klub dan umur rata - rata para pemain sepak bola dalam klub tersebut.
2. Sistem ini membantu bagi penggunanya. Adapun keuntungan yang dapat diambil dengan adanya sistem yang baru ini adalah memprediksi pertandingan sepakbola lebih akurat dengan data - data yang ada, dapat lebih menghemat waktu dalam memprediksi pertandingan sepakbola yang akan bertanding dan mengetahui hasil pertandingan sepakbola tersebut.
3. Algoritma C.45 dengan metode pohon keputusan dapat memberikan informasi prediksi untuk menggambarkan proses yang terkait dengan prediksi klub sepak bola yang meraih kemenangan.

### 4. PENUTUP

Demikianlah hasil penelitian mengenai Penerapan *Data Mining* Dalam Memprediksi Pemenang Klub Sepak Bola Pada Ajang Liga Champion Dengan Algoritma C.45 yang telah dibuat. Penulis berharap agar sistem ini bisa bermanfaat bagi para penggemar sepak bola di dunia sehingga dengan adanya sistem ini pengguna diharapkan tidak menyalahgunakan program ini untuk kepentingan yang tidak - tidak, seiring berkembangnya sepakbola modern maka akan terjadi pula peningkatan sistem sehingga menyebabkan sistem yang ada akan selalu diperbaharui sesuai dengan data - data yang ada maka dari itu pada saatnya nanti perlu dilakukan evaluasi dan pengembangan sistem lebih lanjut dan pengguna perlu melakukan peninjauan terhadap sistem yang diusulkan agar dapat mempertimbangkan baik dan buruknya sistem

sehingga dapat berfungsi sebagaimana yang diharapkan.

### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Prima Indonesia Medan, Dosen Pembimbing, yang telah memberikan kesempatan dan dukungan sehingga penelitian ini dapat terselesaikan dengan baik.

Intan Bandar Lampung)”, Jurnal TIM Darmajaya Vol. 1, No. 2, Oktober 2015, ISSN: 2442-5567.

- [12] Evicienna, Hilda Amalia, “Algoritma C.45 Untuk Prediksi Hasil Pemilihan Legislatif DPRD DKI Jakarta”, Techno Nusa Mandiri Vol. IX, No. 1, Maret 2013.

### DAFTAR PUSTAKA

- [1] Arif Alfa Risal, Abdullah, “Implementasi Metode *Data Mining* Untuk Memprediksi Pertandingan Sepak Bola”, Jurnal Sistemasi Vol. 5, No. 2, Mei 2016, ISSN: 2032-8149.
- [2] Arnold Arbowo, Samuel Lukas, Yast William, “Prediksi Hasil Pertandingan Sepak Bola Menggunakan Jaringan Saraf Tiruan Dan Logika *Fuzzy*”, Konferensi Nasional Sistem dan Informatika 2008.
- [3] Marwana, “Algoritma C.45 Untuk Simulasi Prediksi Kemenangan Dalam Pertandingan Sepak Bola”.
- [4] Wikipedia, “[https://id.wikipedia.org/wiki/Liga\\_Champions\\_UEFA](https://id.wikipedia.org/wiki/Liga_Champions_UEFA)”, 23 Desember 2018 13.11
- [5] Daniel Walangare, Rosa Delima, Restyandito, “Sistem Prediksi Pertandingan Sepak Bola Dengan Metode AHP”.
- [6] Rengga Ryandah Zudha Prawira, Thahja Tribinuka, “Pembinaan Pemain Muda Melalui Akademi Sepak Bola”, Jurnal Sains Dan Seni ITS Vol. 5, No. 2, 2016, ISSN: 2337-3520.
- [7] Gusti Made Trisetia Putra, Muhammad Rusli, “*Arrangement Of Players Position In Soccer Using The Technique Of Naive Bayes*”, Comtech Vol. 6, No. 4, Desember 2015, ISSN: 627-639.
- [8] Nurul Azwanti, “Analisa Algoritma C.45 Untuk Memprediksi Penjualan Motor Pada PT. Capella Dinamik Nusantara Cabang Muka Kuning”, Jurnal Ilmu Ilmiah Komputer Vol. 13, No. 1, Februari 2018, ISSN: 1858-4853.
- [9] Erlin Elisa, “Analisa dan Penerapan Algoritma C.45 Dalam *Data Mining* Untuk Mengidentifikasi Faktor – Faktor Penyebab Kecelakaan Kerja Kontruksi PT. Arupadhatu Adisesanti”, JOIN Vol. 2, No. 1, Juni 2017, ISSN: 2527-9165.
- [10] Fandy Ferdian Harryanto, Seng Hansun, “Penerapan Algoritma C.45 Untuk Memprediksi Penerimaan Calon Pegawai Baru di PT. WISE”, Jatasi Vol. 3, No. 2, Maret 2017.
- [11] Muhammad Arif Rahman, “Algoritma C.45 Untuk Menentukan Mahasiswa Penerima Beasiswa (Studi Kasus : PPS IAIN Raden