



# Plagiarism Checker X Originality Report

**Similarity Found: 7%**

Date: Monday, October 21, 2019

Statistics: 260 words Plagiarized / 3815 Total words

Remarks: Low Plagiarism Detected - Your Document needs Optional Improvement.

---

**PENERAPAN METODE RANDOM FOREST UNTUK PREDIKSI WIN RATIO PEMAIN PLAYER UNKNOWN BATTLEGROUND** 1) Reinardus Aji Haristu 1Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Sanata Dharma Paingan, Maguwoharjo, Depok, Sleman, D.I.Y, Indonesia Telp. (0274) 883037, 889368 Fax (0274) 886529 E-mail: reinardusajih@gmail.com, 2) Paulina H. Prima Rosa 2Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Sanata Dharma Paingan, Maguwoharjo, Depok, Sleman, D.I.Y, Indonesia Telp. (0274) 883037, 889368 Fax (0274) 886529 E-mail: rosa@usd.ac.id ABSTRAKS Game online yang sedang naik daun sebagai salah satu e-sport adalah Player Unknown Battleground.

Untuk memenangkan game tersebut dibutuhkan strategi yang tepat. Dalam makalah ini diuraikan penelitian tentang strategi bermain yang efektif pada game Player Unknown Battleground dengan melakukan penambangan data terhadap data statistik pemain yang diambil dari website kaggle.

Dari data statistik tersebut dibuat model untuk memprediksi win ratio dari setiap pemain menggunakan metode Random Forest. Dari hasil penelitian, Random Forest mampu menghasilkan akurasi sebesar 88,19%. Keyword: Penambangan Data, Klasifikasi, Metode Random Forest, Player Unknown Battleground, PUBG

PENDAHULUAN Perkembangan teknologi di zaman sekarang ini sudah maju sangat pesat. Berbagai inovasi dan terobosan dibidang multimedia/hiburan diciptakan.

Game adalah salah satu inovasi yang tercipta dari efek perkembangan teknologi pada bidang multimedia/hiburan. Saat ini berbagai game online mulai bermunculan, game online merupakan permainan yang dapat dimainkan secara daring (online). Saat ini game sudah bukan lagi hiburan untuk membuang waktu, melainkan sudah berubah menjadi industri baru yang potensial untuk menghasilkan pendapatan yang menggiurkan (Jannah, 2018). E-sport merupakan permainan video game yang kompetitif.

saat ini e-sport begitu banyak dilirik oleh berbagai pihak, dari televisi sampai perusahaan-perusahaan yang bahkan tidak dalam bidang e-sport. Posisi e-sport sama halnya seperti catur yang sudah termasuk dalam kategori olahraga dan sudah ada perlombaan di berbagai olimpiade (Restika, 2018). Saat ini ada beberapa aliran (genre) e-sport yang terkenal, diantaranya adalah MOBA (Multiplayer Online Battle Arena), FPS (First Person Shooter), Fighting (Street Fighter), Real Time Strategy Games (Budi, 2018). Sudah banyak kompetisi yang diadakan untuk pertandingan e-sport dengan hadiah yang tidak sedikit.

Salah satu kompetisi yang terbaru adalah turnamen e-sport Asia Tenggara yang diselenggarakan oleh pemerintah Indonesia dengan total hadiah mencapai Rp1,4 M dan dilaksanakan di Jakarta pada tanggal 17 Oktober sampai 21 Oktober 2018. Game yang dipertandingkan antara lain adalah Mobile Legend, Arena Of Valor (AOV), Playerunknown's Battleground, Point Blank dan DOTA 2.

Jika dilihat dari penyelenggara dan dari nominal hadiah yang diperebutkan, maka e-sport tidak bisa dipandang rendah. Potensi industri dalam bidang tersebut memang ada dan sudah terlihat. Salah satu game yang termasuk dalam golongan e-sport yang sedang naik daun yaitu Player Unknown Battleground atau yang sering disebut dengan PUBG.

Game ini termasuk dalam genre permainan FPS (First Person Shooter) namun juga dilengkapi dengan mode Third Person. Permainan ini memiliki beberapa mode bermain berdasar kategori keanggotaan yaitu solo (sendirian/tanpa tim), duo (tim beranggotakan dua orang) dan squad (tim beranggotakan empat orang). Konsep permainan Player Unknown Battleground menggunakan konsep survival, dimana pemain yang bertahan hidup terakhir maka yang akan menang.

Selama pertandingan pemain diminta untuk mencari segala senjata dan perlengkapan

yang dibutuhkan untuk bertahan hidup pada suatu pulau. Dalam permainan ini ada istilah win ratio yaitu nilai persentase yang menunjukkan rasio seberapa besar ia menang selama ia bermain dari awal ia mendaftarkan akun sampai terakhir ia bermain. Tidak mudah untuk bertahan hidup dan menjadi yang terakhir dalam pertandingan game tersebut.

Terkadang butuh strategi yang matang untuk memenangkan pertandingan. Selama ini belum ada penelitian terkait pencarian strategi terbaik dalam bermain game ini. Dalam penelitian ini penulis akan menganalisis rekaman data statistik dari berbagai pemain dengan tujuan untuk mencari strategi terbaik dalam bermain game PUBG. Ada beberapa penelitian yang sudah dilakukan dengan menggunakan metode Random Forest. Salah satunya adalah yang dilakukan oleh Nidhomiddin dan Otok (2015).

Dalam penelitiannya mereka melakukan klasifikasi penderita HIV/AIDS menggunakan metode Random Forest dan Multivariate Adaptive Regression Spline (MARS) binary response. Hasil yang didapat adalah mengetahui variabel yang paling berpengaruh untuk menentukan status HIV/AIDS yaitu usia, jenis pekerjaan, pernah ditahan karena kasus NAPZA, status nikah dan selalu pakai jarum steril. Akurasi yang didapatkan ketika menggunakan metode MARS sebesar 80,28%.

Sedangkan ketika menggunakan metode Random Forest diperoleh akurasi terbaik yaitu 97,80%. Penulis tersebut juga mencoba gabungan antara metode MARS dan Random Forest, dan memperoleh hasil akurasi sebesar 91,00%. Penelitian lain adalah yang dilakukan oleh Nugroho dan Emiliyawati (2017).

Dalam penelitiannya mereka melakukan klasifikasi tingkat penerimaan konsumen terhadap mobil menggunakan metode Random Forest. Dari hasil penelitian, mereka mendapatkan bahwa variabel yang mempengaruhi tingkat penerimaan konsumen terdiri dari harga pembelian, biaya perawatan, jumlah pintu, kapasitas penumpang, ukuran bagasi dan taksiran keselamatan penumpang.

Berdasar latar belakang tersebut, dalam makalah ini penulis menggunakan metode Random Forest untuk memprediksi win ratio terhadap data statistik pemain game Player Unknown Battleground dengan menggunakan metode Random Forest. Dengan demikian diharapkan dapat dihasilkan model untuk memprediksi win ratio. Model tersebut memiliki beberapa tree yang setiap node merupakan atribut yang dipilih karena kemampuannya yang baik untuk mengklasifikasi.

Atribut yang dijadikan Node pada tree dapat dipertimbangkan sebagai strategi dalam bermain game PUBG. Penelitian ini akan mencoba mengetahui bagaimana menerapkan

metode Random Forest untuk memprediksi nilai win ratio dengan baik berikut atribut (fitur) apa yang paling berpengaruh dalam memprediksi nilai win ratio serta mengetahui berapa akurasi prediksi terbaik yang dihasilkan oleh metode Random Forest terhadap data statistik pemain game Player Unknown Battleground.

METODOLOGI Berikut ini adalah tahap-tahap penelitian yang dilakukan oleh penulis: Pengumpulan Data Tahap awal dalam penelitian ini adalah pengumpulan data. Pengumpulan data yang dilakukan dalam penelitian ini dengan cara mengunduh data dari website Kaggle.com. Knowledge Discovery in Database (KDD) Dengan mengacu proses KDD menurut Han (2006), berikut ini adalah proses KDD yang dilakukan pada penelitian ini: Data Selection Dalam tahap seleksi data akan dilakukan pemilihan atribut yang relevan untuk digunakan dalam penelitian. Atribut yang tidak digunakan atau tidak relevan akan dihapus.

Atribut yang akan dihapus dari data statistik Pemain Player Unknown Battleground adalah atribut yang tidak independen. Berikut ini adalah daftar atribut yang dihapus beserta alasannya yang ditunjukkan pada Tabel 1. Dengan demikian didapatkan atribut yang akan digunakan pada Tabel 2.

Sebagai catatan, atribut solo\_RoundsPlayed akan disimpan untuk pengujian perangkat lunak menggunakan dataset. Tabel 1 Daftar atribut yang dihapus

No	Nama Atribut	Alasan Penghapusan
1	player_name	Memiliki nilai unik (berbeda pada setiap baris data)
2	tracker_id	Memiliki nilai unik (berbeda pada setiap baris data)
3	solo_RoundsPlayed	Digunakan untuk menghitung atribut solo_WinRatio
4	solo_Wins	Digunakan untuk menghitung atribut solo_WinRatio
5	solo_TimeSurvived	Digunakan untuk menghitung atribut solo_TimeSurvivedPg
6	solo_Top10s	Digunakan untuk menghitung atribut solo_WinTop10Ratio
7	solo_Kills	Digunakan untuk menghitung atribut solo_killDeathRatio
8	solo_Assists	Tidak memiliki nilai untuk permainan mode solo
9	solo_TeamKills	Tidak memiliki arti untuk permainan mode solo
10	solo_TeamKillsPg	Tidak memiliki arti untuk permainan mode solo
11	solo_HeadshotKills	Digunakan untuk menghitung atribut solo_HeadshotKillsPg
12	solo_RoadKills	Digunakan untuk menghitung atribut solo_RoadKillsPg
13	solo_Days	Digunakan untuk menghitung atribut solo_DailyKills
14	solo_MostSurvivalTime	Memiliki nilai yang sama dengan atribut solo_LongestTimeSurvived
15	solo_WalkDistance	Digunakan untuk menghitung atribut solo_AvgWalkDistance
16	solo_MoveDistance	Digunakan untuk menghitung atribut solo_RideDistance+AP1
17	solo_RideDistance+AP1	Digunakan untuk menghitung atribut solo_AvgRideDistance
18	solo_Revives	Tidak memiliki arti untuk permainan mode solo (menyelamatkan teman satu tim)
19	solo_WeaponAcquired	Memiliki nilai nol untuk semua baris data
20	solo_Assists	Tidak memiliki arti untuk permainan mode

solo (membantu teman satu tim) \_21\_solo\_DBNOs\_Memiliki nilai nol untuk semua baris data \_22\_Solo\_WinTop10Ratio\_Menggunakan atribut solo\_WinRatio (atribut label/kelas) untuk menghitung nilai \_Tabel 2 Daftar atribut yang dihapus No\_Nama Atribut\_Keterangan \_1\_solo\_WinRatio\_Rasio kemenangan \_2\_solo\_killDeathRatio\_Ratio jumlah korban \_3\_solo\_Top10Ratio\_Rasio kemenangan dalam 10 besar \_4\_solo\_Losses\_Jumlah kekalahan \_5\_solo\_Rating\_Rating yang dimiliki terakhir kali \_6\_solo\_BestRating\_Rating terbaik selama bermain \_7\_solo\_DamagePg\_Jumlah damage yang diciptakan per ronde \_8\_solo\_HeadshotKillsPg\_Jumlah head shot yang diciptakan per ronde \_9\_solo\_HealsPg\_Jumlah poin heal (penyembuhan) per ronde \_10\_solo\_KillsPg\_Jumlah korban per ronde \_11\_solo\_MoveDistancePg\_Jarak yang ditempuh per ronde \_12\_solo\_RoadKillsPg\_Rata-rata korban yang dibunuh menggunakan kendaraan \_13\_solo\_TimeSurvivedPg\_Waktu bertahan hidup per ronde \_14\_solo\_Top10sPg\_Rata-rata kemenangan di sepuluh besar \_15\_solo\_Suicides\_Jumlah kejadian bunuh diri (mati tanpa diserang musuh) \_16\_solo\_HeadshotKillRatio\_Rasio head shot (korban tepat kepala) \_17\_solo\_VehicleDestroys\_Jumlah penghancuran kendaraan \_18\_solo\_DailyKills\_Jumlah korban harian (rata-rata korban dalam satu hari) \_19\_solo\_WeeklyKills\_Jumlah korban minggu (rata-rata korban dalam satu minggu) \_20\_solo\_MaxKillStreaks\_Jumlah korban terbanyak yang diciptakan berturut-turut \_21\_solo\_LongestTimeSurvived\_Waktu terlama bertahan hidup \_22\_solo\_AvgSurvivalTime\_Rata-rata waktu bertahan hidup \_23\_solo\_WinPoints\_Jumlah poin kemenangan terakhir \_24\_solo\_AvgWalkDistance\_Rata-rata jarak yang sudah ditempuh dengan jalan kaki \_25\_solo\_AvgRideDistance\_Rata-rata jarak yang sudah ditempuh dengan kendaraan \_26\_solo\_LongestKill\_Jarak korban terjauh \_27\_solo\_Heals\_Jumlah poin heal (penyembuhan) \_28\_solo\_Boosts\_Total energi yang digunakan \_29\_solo\_DamageDealt\_Jumlah poin damage (kerusakan) yang diterima \_30\_solo\_RoundMostKills\_Jumlah korban terbanyak dalam satu ronde \_ Data transformation Ada beberapa transformasi data yang dilakukan penulis pada dataset statistik **Pemain Player Unknown Battleground.**

Transformasi yang pertama adalah merubah nilai atribut solo\_WinRatio dari nilai yang kontinu menjadi kategorial dengan cara mengelompokkan nilai tersebut menjadi tiga kelompok. Dalam mengelompokkan nilai atribut solo\_WinRatio, penulis menggunakan tools Weka versi 3.8. Metode yang penulis pilih adalah metode simple k means dengan hasil nilai centroid akhir atribut solo\_WinRatio yang ditunjukkan pada Gambar 1.

Selain atribut solo\_WinRatio, atribut solo\_RoundsPlayed juga akan dikelompokkan (clustering) menjadi 3 kelompok yang menggambarkan tipe pemain yaitu pemula, normal, profesional. Cara pengelompokannya pun sama seperti mengelompokkan atribut solo\_WinRatio, yaitu menggunakan tools Weka versi 3.8 dengan metode simple k means. Hasil nilai akhir centroid atribut solo\_RoundsPlayed **ditunjukkan pada Gambar 2.**

Pada Gambar 3 merupakan hasil klustering untuk atribut solo\_WinRatio setelah menghilangkan kelompok pemain dengan kluster 0 (pemula) berdasar atribut solo\_RoundsPlayed (Gambar 2). / Gambar 1 Hasil klustering tools weka untuk atribut solo\_WinRatio // Gambar 2 Hasil klustering tools weka untuk atribut solo\_RoundsPlayed / Gambar 3 Hasil Klustering Tools Weka Untuk Atribut solo\_WinRatio Setelah Menghilangkan Kluster 0 atribut solo\_RoundsPlayed Transformasi yang kedua adalah dengan menormalisasi seluruh nilai atribut kedalam range/jarak yang sama yaitu nol sampai satu.

Proses normalisasi dilakukan dengan menggunakan metode min-max normalization, adapun rumus yang digunakan dalam metode tersebut dideskripsikan pada Persamaan 1 berikut: 
$$x' = \frac{x - \min(A)}{\max(A) - \min(A)}$$
 (1) Keterangan:  $x'$  : nilai setelah ternormalisasi  $x$  : nilai sebelum ternormalisasi  $\min(A)$  : nilai minimal dari atribut A  $\max(A)$  : nilai maksimal dari atribut A  $\min(A)$  : nilai minimal terbaru dari atribut A  $\max(A)$  : nilai maksimal terbaru dari atribut A Data mining Proses **data mining adalah proses penggalian** informasi dari data yang ada.

Penggalian informasi pada penelitian ini berupa klasifikasi. Dalam penelitian ini proses klasifikasi akan dilakukan **menggunakan metode Random Forest untuk** data statistik pemain PUBG. Random Forest merupakan sebuah metode ensemble. Metode ensemble merupakan cara untuk meningkatkan akurasi metode klasifikasi dengan cara mengkombinasikan metode klasifikasi (Han, 2012).

Random Forest diawali dengan teknik dasar data mining yaitu decision tree. Pada decision tree input dimasukkan pada bagian atas (root) kemudian turun kebagian bawah (leaf) untuk menentukan data tersebut termasuk kelas apa. **Random forest adalah pengklasifikasi yang terdiri dari kumpulan pengklasifikasi** pohon terstruktur dimana **masing-masing pohon melemparkan unit suara untuk kelas** paling populer di input x (Breiman, 2001).

Dengan kata lain Random Forest terdiri dari sekumpulan decision tree (pohon keputusan), dimana kumpulan decision tree tersebut digunakan untuk mengklasifikasi data ke suatu kelas. Random Forest merupakan sebuah metode ensemble. Metode ensemble merupakan cara untuk meningkatkan akurasi metode klasifikasi dengan cara mengkombinasikan metode klasifikasi (Han, 2012).

Random Forest diawali dengan teknik dasar data mining yaitu decision tree. Pada decision tree input dimasukkan pada bagian atas (root) kemudian turun kebagian bawah

(leaf) untuk menentukan data tersebut termasuk kelas apa. Random forest adalah pengklasifikasi yang terdiri dari kumpulan pengklasifikasi pohon terstruktur dimana masing-masing pohon melemparkan unit suara untuk kelas paling populer di input  $x$  (Breiman, 2001).

Dengan kata lain Random Forest terdiri dari sekumpulan decision tree (pohon keputusan), dimana kumpulan decision tree tersebut digunakan untuk mengklasifikasi data ke suatu kelas. Random forest merupakan metode klasifikasi yang supervised. Sesuai dengan namanya, metode ini menciptakan sebuah hutan (forest) dengan sejumlah pohon (tree).

Secara umum, semakin banyak pohon (tree) pada sebuah hutan (forest) maka semakin kuat juga hutan tersebut terlihat. Pada kasus yang sama, semakin banyak tree, maka semakin besar pula akurasi yang didapatkan (Polamuri, 2017). Decision Tree akan menggunakan information gain untuk perhitungan dalam menentukan root node dan rule.

Sama halnya dengan Random Forest yang akan menggunakan information gain dan gini index untuk perhitungan dalam membangun tree (Han, 2012), hanya saja Random Forest akan membangun lebih dari satu tree. Masing-masing tree dibangun menggunakan set data dengan atribut yang diambil secara acak dari data training. Dengan kata lain setiap tree akan bergantung pada nilai dari sampel vektor yang independen dengan distribusi yang sama pada setiap tree (Han, 2012).

Selama proses klasifikasi setiap tree akan memberikan voting kelas yang paling populer (Han, 2012). Berikut ini adalah algoritma untuk metode random forest. Algoritma dibagi menjadi dua bagian, bagian pertama adalah pembuatan " $n$ " pohon (tree) untuk membentuk hutan (forest) yang acak (random). Bagian kedua adalah algoritma untuk melakukan prediksi dari Random Forest yang sudah dibuat (Han, 2012).

Input:  $D$ , dataset yang terdiri dari  $d$  baris  $k$ , angka dari jumlah tree  
Algoritma untuk membentuk model Random Forest (Bagian 1):  
1. Buat data sampel data  $D_i$  dengan mengambil acak dari dataset  $D$  dengan pengembalian (replacement)  
2. Gunakan sampel data  $D_i$  untuk membangun tree ke  $i$  ( $i = 1, 2, \dots, k$ )  
3. Ulangi langkah a dan b sebanyak  $k$   
Langkah metode Random Forest diawali dengan pemilihan " $k$ " sampel dataset  $D_i$  yang diambil secara acak dengan pengembalian (replacement).

Langkah selanjutnya adalah menggunakan dataset  $D_i$  untuk membangun decision tree ke- $i$ . Dalam membangun tree ke- $i$ , metodologi CART dapat digunakan. Metodologi CART menggunakan information gain dalam menentukan setiap node pada tree.

Perhitungan information gain dapat dihitung menggunakan Persamaan 2.  $IG(A) = H(D) - \sum_{i=1}^k P_i H(D_i)$  (2) Dimana nilai  $IG(A)$  dapat dicari menggunakan Persamaan 3 dan  $H(D)$  dicari menggunakan Persamaan 4 dibawah ini:  $H(D) = - \sum_{i=1}^k P_i \log_2 P_i$  (3) Keterangan:  $k$  : jumlah kelas target  $P_i$  : probabilitas munculnya kelas ke  $i$  pada partisi  $D$   $H(D_i) = - \sum_{j=1}^k P_{ij} \log_2 P_{ij}$  (4) Keterangan:  $k$  : jumlah partisi  $P_{ij}$  : total partisi ke  $j$   $n_j$  : jumlah tuple/baris pada semua partisi Untuk menghitung information gain dari atribut yang memiliki nilai kontinu (numerik), maka harus ditentukan nilai pembelah (split-point) terbaik untuk mengelompokan nilai dari atribut tersebut. Untuk mencari split-point terbaik maka data dari atribut tersebut harus diurutkan terlebih dahulu.

Nilai tengah antara setiap pasangan nilai yang berdekatan dianggap sebagai kemungkinan yang bisa dijadikan split-point (Han, 2012). Pseudocode untuk proses prediksi data test (Bagian 2): Ambil data test dan gunakan rule dari setiap tree untuk memprediksi keluaran klasifikasi dari data tersebut, simpan hasil yang didapat Hitung suara (vote) untuk setiap target yang diprediksi dari setiap tree Pertimbangkan target prediksi yang terpilih dengan memilih target kelas yang paling banyak diprediksi sebagai hasil prediksi akhir dari metode random forest Untuk memprediksi kelas target dari data test menggunakan random forest, masukan data test melalui aturan-aturan (rule) yang sudah dibuat menggunakan tree.

Hasil prediksi setiap tree bisa saja ada yang berbeda dan ada yang sama, maka prediksi akhir akan dipilih berdasar prediksi kelas yang terbanyak diprediksi. Misalkan dari 100 tree, 80 tree memprediksi target adalah kelas A dan sisanya adalah kelas B, maka prediksi akhir yang dipilih adalah kelas A konsep pemilihan dengan suara terbanyak yaitu hasil prediksi dari semua tree dinamakan majority voting.

Pengujian yang akan dilakukan, penulis menggunakan empat jenis dataset: Dataset statistik Pemain Player Unknown Battleground dengan outlier dan data tidak ternormalisasi. Dataset statistik Pemain Player Unknown Battleground dengan outlier dan data ternormalisasi. Dataset statistik Pemain Player Unknown Battleground tanpa outlier dan data tidak ternormalisasi.

Dataset statistik Pemain Player Unknown Battleground tanpa outlier dan data ternormalisasi. Pada pengujian tersebut akan dilakukan pengujian klasifikasi dengan jumlah fold 3 dan kombinasi jumlah tree yang berbeda-beda dimulai dari 10 sampai 90 dengan interval 10. Dari pengujian ini akan dilihat hasil akurasi dari masing-masing pengujian.

Pattern evaluation dan Knowledge presentation Untuk proses mengidentifikasi pola

yang tepat yang merupakan hasil dari proses data mining (Pattern evaluation) dan proses penyajian hasil dari data mining kepada user (Knowledge presentation) dilakukan setelah sistem selesai dibangun dan proses penambahan data selesai dilakukan. Dalam proses ini penulis melakukan evaluasi dari hasil penambahan data yang didapat dari perangkat lunak yang telah dibangun dan menjelaskan hasil evaluasi tersebut agar informasi yang didapat dapat dengan mudah diterima oleh pihak-pihak yang membutuhkan. Kedua langkah tersebut akan dipaparkan pada bagian pembahasan.

Pengembangan Perangkat Lunak Penulis melakukan pengembangan perangkat lunak sebagai alat untuk mengolah dataset yang dimiliki penulis untuk mendapatkan informasi (knowledge) yang berguna. Metode yang digunakan oleh penulis adalah metode waterfall. Metode tersebut adalah salah satu metode yang sangat sering digunakan oleh para pengembang perangkat lunak.

Pengerjaan sistem secara linear diberlakukan dalam metode ini. Dimana jika tahap pertama belum selesai maka tahap kedua belum bisa dilakukan. Analisis Hasil Dalam tahap ini dilakukan analisa kinerja terhadap model atau metode Random Forest yang sudah diimplementasikan ke dalam sebuah perangkat lunak. PEMBAHASAN Hasil Pengujian Dataset Berikut adalah hasil dari pengujian empat dataset: Pengujian Dataset statistik Pemain Player Unknown Battleground dengan outlier dan data tidak ternormalisasi Dataset statistik Pemain Player Unknown Battleground dengan outlier dan data tidak ternormalisasi memiliki 87.898 baris data dan 30 atribut. Tabel pengujian dataset statistik Pemain Player Unknown Battleground tanpa outlier dapat dilihat di Tabel 3.

Tabel 3 Pengujian dataset statistik Pemain Player Unknown Battleground dengan outlier dan data tidak ternormalisasi No \_Jumlah Tree \_Akurasi (%) \_1 \_10 \_86,11 \_2 \_20 \_87,14 \_3 \_30 \_87,55 \_4 \_40 \_87,85 \_5 \_50 \_87,85 \_6 \_60 \_87,86 \_7 \_70 \_88,19 \_8 \_80 \_87,81 \_ \_ Pengujian Dataset statistik Pemain Player Unknown Battleground dengan outlier dan data ternormalisasi Dataset statistik Pemain Player Unknown Battleground dengan outlier dan data ternormalisasi memiliki 87.898 baris data dan 30 atribut. Tabel pengujian Dataset statistik Pemain Player Unknown Battleground dengan outlier dan data ternormalisasi dapat dilihat di Tabel 4.

Tabel 4 Pengujian dataset statistik Pemain Player Unknown Battleground dengan outlier dan data ternormalisasi No \_Jumlah Tree \_Akurasi (%) \_1 \_10 \_86,11 \_2 \_20 \_87,14 \_3 \_30 \_87,55 \_4 \_40 \_87,85 \_5 \_50 \_87,85 \_6 \_60 \_87,86 \_7 \_70 \_88,19 \_8 \_80 \_87,81 \_ \_ Pengujian Dataset statistik Pemain Player Unknown Battleground tanpa outlier dan data tidak ternormalisasi Dataset statistik Pemain Player Unknown Battleground tanpa outlier dan data tidak ternormalisasi memiliki 58.366 baris data dan 30 atribut.

Tabel pengujian dataset statistik **Pemain Player Unknown Battleground** tanpa outlier dan data tidak ternormalisasi dapat dilihat di Tabel 5. Tabel 5 Pengujian dataset statistik **Pemain Player Unknown Battleground** tanpa outlier dan data tidak ternormalisasi No \_Jumlah Tree \_Akurasi (%) \_1 \_10 \_86,11 \_2 \_20 \_87,14 \_3 \_30 \_87,55 \_4 \_40 \_87,85 \_5 \_50 \_87,85 \_6 \_60 \_87,86 \_7 \_70 \_88,19 \_8 \_80 \_87,81 \_ \_ Pengujian Dataset statistik **Pemain Player Unknown Battleground** tanpa outlier dan data ternormalisasi Dataset statistik **Pemain Player Unknown Battleground** tanpa outlier dan data ternormalisasi memiliki 58.366 baris data dan 30 atribut. Tabel pengujian dataset statistik **Pemain Player Unknown Battleground** dengan outlier dan data ternormalisasi dapat dilihat di Tabel 6.

Tabel 6 Pengujian dataset statistik **Pemain Player Unknown Battleground** dengan outlier dan data ternormalisasi No \_Jumlah Tree \_Akurasi (%) \_1 \_10 \_86,11 \_2 \_20 \_87,14 \_3 \_30 \_87,55 \_4 \_40 \_87,85 \_5 \_50 \_87,85 \_6 \_60 \_87,86 \_7 \_70 \_88,19 \_8 \_80 \_87,81 \_ \_ Evaluasi Hasil Pengujian Dari hasil pengujian ke empat dataset tersebut terdapat kesamaan pola pada pengujian keempat dataset tersebut yaitu akurasi yang cenderung bertambah ketika jumlah tree semakin besar.

Pola tersebut dapat mudah di lihat pada grafik yang digambarkan pada Gambar 4 sampai 7. Hal tersebut menunjukkan bahwa semakin banyak jumlah tree yang digunakan dalam membangun model maka akurasi cenderung meningkat. / Gambar 4 Hasil pengujian dataset statistik **Pemain Player Unknown Battleground** dengan outlier dan data tidak ternormalisasi / Gambar 5 Hasil pengujian dataset statistik **Pemain Player Unknown Battleground** dengan outlier dan data ternormalisasi / Gambar 6 Hasil pengujian dataset statistik **Pemain Player Unknown Battleground** tanpa outlier dan data tidak ternormalisasi / Gambar 7 Hasil pengujian dataset statistik **Pemain Player Unknown Battleground** tanpa outlier dan data ternormalisasi Dari hasil pengujian keempat dataset tersebut didapatkan akurasi tertinggi yaitu 88,19%.

Akurasi tersebut didapat dari hasil klasifikasi dataset statistik **Pemain Player Unknown Battleground** dengan outlier dan data tidak ternormalisasi dengan jumlah tree sebanyak 70. Jika dilihat model yang dibuat untuk mendapatkan hasil tertinggi tersebut, atribut yang sering digunakan sebagai root node maupun node lain adalah atribut solo\_KillDeathRatio.

Artinya atribut tersebut sangat berperan besar dalam memprediksi atau mengklasifikasi win ratio. Dengan memanfaatkan tools Weka, penulis mencoba untuk melihat bagaimana korelasi atribut tersebut dengan win ratio. Jika dilihat dari scatter plot pada Gambar 8, ada cukup banyak nilai solo\_KillDeathRatio yang tinggi yang menunjukkan win

ratio tinggi (kluster 2) artinya jika nilai split diletakan pada tengah atribut solo\_KillDeathRatio maka ada cukup banyak data yang memiliki win ratio tinggi (kluster 2) yang terdapat pada salah satu subset.

Hal tersebutlah yang menyebabkan atribut ini selalu memiliki nilai information gain tinggi dari pada atribut lain. / Gambar 8 scatter plot atribut solo\_KillDeathRatio dengan win ratio KESIMPULAN Hasil penelitian **penerapan metode random forest untuk prediksi win ratio pemain player unknown battleground** ini menghasilkan kesimpulan bahwa Metode random forest dapat digunakan untuk melakukan prediksi/klasifikasi **win ratio pemain player unknown battleground** dengan baik. Semakin banyak jumlah tree yang digunakan maka semakin baik pula akurasi yang didapat.

Atribut yang paling berpengaruh dalam menentukan klasifikasi **win ratio pemain player unknown battleground** adalah solo\_KillDeathRatio. Untuk Akurasi prediksi terbaik yang dihasilkan oleh metode Random Forest terhadap win ratio pada data statistik pemain game Player Unknown Battleground adalah sebesar 88,19%.

SARAN Untuk pengembangan penelitian yang akan datang dapat ditambahkan fungsi untuk menyimpan atau mengekspor model yang berhasil dibangun oleh perangkat lunak. **Selain itu dapat juga** ditambahkan fungsi untuk menampilkan semua rule yang memiliki leaf suatu label. Untuk mempermudah pemrosesan awal (preprocessing) juga dapat ditambahkan fungsi untuk melakukan transformasi/normalisasi data dengan secara otomatis dari sistem.

Pada penelitian selanjutnya juga bisa dicobakan dataset yang berbeda atau metode yang berbeda. PUSTAKA Han, Jiawei.2012. Data Mining Concepts and Techniques Third Edition. USA:Elsevier IncJati, Anggoro Suryo.2018.UniPin Bikin Kompetisi eSports Berhadiah Rp 1,4 Miliar di <https://inet.detik.com/games-news/d-4162385/unipin-bikin-kompetisi-esports-berhadiah-rp-14-miliar> (di akses November) Jannah, Selfie Miftahul.2018. Bukan Cuma Main Game, Esport Mulai Jadi Industri Masa Depan di <https://finance.detik.com/berita-ekonomi-bisnis/d-4316768/bukan-cuma-main-game-esport-mulai-jadi-industri-masa-depan> (di akses November) Nidhomiddin & Otok.2015.RANDOM **FOREST DAN MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS) BINARY RESPONSE UNTUK KLASIFIKASI PENDERITA HIV/AIDS DI SURABAYA**.Bandung:Institut Teknologi Sepuluh November Surabaya. Nugroho.2017.Sistem **Klasifikasi Variabel Tingkat Penerimaan Konsumen Terhadap Mobil Menggunakan Metode Random Forest** di [https://www.researchgate.net/publication/320413581\\_Sistem\\_Klasifikasi\\_Variabel\\_Tingkat\\_Penerimaan\\_Konsumen\\_Terdapat\\_Mobil\\_Menggunakan\\_Metode\\_Random\\_Forest](https://www.researchgate.net/publication/320413581_Sistem_Klasifikasi_Variabel_Tingkat_Penerimaan_Konsumen_Terdapat_Mobil_Menggunakan_Metode_Random_Forest)

(diakses November) Polamuri, Saimadhu. How **Random Forest Algorithm Works In Machine Learning**

<https://dataaspirant.com/2017/05/22/random-forest-algorithm-machine-learning>

Refaeilzadeh, Payam.2008.Cross Validation di

<http://leitang.net/papers/ency-cross-validation.pdf> (diakses November) Restika,

Ria.2018. Apa Itu Esports? di <https://esportsnesia.com/penting/apa-itu-esports/> (di akses November)

## INTERNET SOURCES:

---

1% - <http://repository.usd.ac.id/view/subjects/Inf=5Feng.html>  
<1% - <https://www.posciety.com/game-pubg-haram-mui/>  
<1% - <http://wintexaspoker.com/>  
<1% - [https://id.wikipedia.org/wiki/Permainan\\_online](https://id.wikipedia.org/wiki/Permainan_online)  
<1% -  
<https://denotnote.blogspot.com/2017/03/pengantar-teknologi-game-tugas-1.html>  
<1% - <https://caberawit57.blogspot.com/>  
<1% -  
<https://mediaformasi.com/2018/10/keren-ini-dia-profil-guest-star-seaca-dan-wesg-2018/>  
<1% - [https://issuu.com/redaksiinfoplus/docs/pdf\\_infoplus\\_04\\_2019](https://issuu.com/redaksiinfoplus/docs/pdf_infoplus_04_2019)  
<1% -  
[https://www.kompasiana.com/femiardine/mengenal-tipetipe-game\\_552a43b2f17e61b46fd6243b](https://www.kompasiana.com/femiardine/mengenal-tipetipe-game_552a43b2f17e61b46fd6243b)  
<1% - <https://siddix.blogspot.com/2018/12/tips-bertahan-hidup-pada-game-pubg.html>  
<1% -  
[https://www.academia.edu/34867977/Sistem\\_Klasifikasi\\_Variabel\\_Tingkat\\_Penerimaan\\_Konsumen\\_Terhadap\\_Mobil\\_Menggunakan\\_Metode\\_Random\\_Forest](https://www.academia.edu/34867977/Sistem_Klasifikasi_Variabel_Tingkat_Penerimaan_Konsumen_Terhadap_Mobil_Menggunakan_Metode_Random_Forest)  
1% -  
[https://www.researchgate.net/publication/320413581\\_Sistem\\_Klasifikasi\\_Variabel\\_Tingkat\\_Penerimaan\\_Konsumen\\_Terhadap\\_Mobil\\_Menggunakan\\_Metode\\_Random\\_Forest](https://www.researchgate.net/publication/320413581_Sistem_Klasifikasi_Variabel_Tingkat_Penerimaan_Konsumen_Terhadap_Mobil_Menggunakan_Metode_Random_Forest)  
<1% - <http://eprints.umm.ac.id/51848/2/BAB%20I.pdf>  
1% - <https://journal.unnes.ac.id/nju/index.php/jte/article/download/10452/6660>  
<1% -  
<http://library.binus.ac.id/eColls/eThesisdoc/Bab2/2012-1-00314-SI%20Bab2001.pdf>  
<1% - <https://ejournal.bsi.ac.id/ejurnal/index.php/evolusi/article/download/604/495>  
<1% - <https://katakamus.id/katakamus/daftaristilah/ekonomi>  
<1% -  
<https://www.dapodik.co.id/2019/08/panduan-lengkap-pengisian-sarpras-tabel.html>  
<1% -  
<https://id.123dok.com/document/8yd076zp-algoritma-modified-k-means-clustering-pada-penentuan-cluster-centre-berbasis-sum-of-squared-error-sse.html>  
<1% - <https://artikel50.blogspot.com/2010/06/normal-0-false-false-false.html>  
1% -  
[https://www.researchgate.net/publication/326345102\\_KOMPARASI\\_ALGORITMA\\_KLASIFIKASI\\_MENENTUKAN\\_KELULUSAN\\_MATA\\_KULIAH\\_PADA\\_UNIVERSITAS](https://www.researchgate.net/publication/326345102_KOMPARASI_ALGORITMA_KLASIFIKASI_MENENTUKAN_KELULUSAN_MATA_KULIAH_PADA_UNIVERSITAS)

<1% - <http://ejurnal.its.ac.id/index.php/teknik/article/download/21120/3463>  
<1% -  
<https://repository.ipb.ac.id/bitstream/handle/123456789/95209/G18ado.pdf?sequence=1&isAllowed=y>  
<1% - <https://endroandriyanto.blogspot.com/2012/06/induksi-decision-tree.html>  
<1% -  
[http://sinasinderaja.lapan.go.id/files/sinasja2014/prosiding/bukuprosiding\\_120-127.pdf](http://sinasinderaja.lapan.go.id/files/sinasja2014/prosiding/bukuprosiding_120-127.pdf)  
<1% -  
[https://www.researchgate.net/publication/294085772\\_Klasifikasi\\_dan\\_Klastering\\_Penjuru\\_san\\_Siswa\\_SMA\\_Negeri\\_3\\_Boyolali](https://www.researchgate.net/publication/294085772_Klasifikasi_dan_Klastering_Penjuru_san_Siswa_SMA_Negeri_3_Boyolali)  
<1% - <https://nichonotes.blogspot.com/2014/11/penyusutan-aset-tetap.html>  
<1% -  
<https://naynaedelweis.blogspot.com/2013/01/tugas-laporan-kelompok-spk-ahp.html>  
<1% - <https://keamanan-informasi.stei.itb.ac.id/>  
<1% - <http://dataaspirant.com/2017/05/22/random-forest-algorithm-machine-learning/>