

# Membangun Pythagoras Sebagai Visualisasi Random Forest Untuk Pemodelan Pohon Keputusan

Erlin Windia Ambarsari<sup>1</sup>, Herlinda<sup>2</sup>

<sup>1,2</sup>Department of Informatic, Universitas Indraprasta PGRI, Indonesia

---

## Article Info

### Article history:

Received June 24, 2020

Revised Oct 14, 2020

Accepted Oct 26, 2020

---

### Keywords:

Pythagoras  
Decision Tree  
Random Forest

---

## ABSTRACT

Students observed Pythagoras for using a plane Geometry and 3D Geometry. However, Pythagoras can also be built for decision trees. Our research regarding Instagram Usage Habit with construct Pythagoras for a single decision tree. The study's results obtained are ambiguous attribute values. Therefore, it is continued with research to build Pythagoras for Random Forest. The purpose of the study is to facilitate the tracking of ambiguous data contained in the attributes. The results obtained that the relationship between characteristics of the target class, thus resulting in misclassification. This error caused invalid data; for example, there are three times the separation of data on the same attribute for age's target for a group of 20. However, although there are misclassifications caused by invalid data, based on the Pythagorean construction for Random Forest, the data is more easily traced to errors, which cannot be done by a single decision tree.

Copyright © 2020 Universitas Indraprasta PGRI.  
All rights reserved.

---

## Corresponding Author:

Erlin Windia Ambarsari,  
Department of Informatic,  
Universitas Indraprasta PGRI,  
Jl. Nangka No. 58 C, Tanjung Barat, Jagakarsa, Jakarta Selatan.  
Email: [erlin.windia@unindra.ac.id](mailto:erlin.windia@unindra.ac.id)

---

## 1. PENDAHULUAN

Dalam pelajaran Matematika, sering diajarkan kepada para peserta didik mengenai masalah teorema, salah satunya teorema Pythagoras. Pythagoras, sesuai dengan nama penemunya, merumuskan sisi miring yang dikuadratkan mempunyai jumlah yang sama dengan keseluruhan total dari kuadrat kedua kaki, dengan membentuk segitiga siku-siku atau hubungan antara kedua kaki untuk membuat sudut 90°. Pythagoras yang sebenarnya sudah lama ditemukan pada zaman Babilonia, mempunyai 371 pembuktian teorema [1].

Pembuktian teorema ini dilakukan pada penelitian dari [2] dalam penguasaan konsep teorema yang diterapkan kepada Siswa kelas VIII MTs Negeri Balang-Balang. Teorema yang diambil pada penelitian tersebut membahas luas daerah persegi dari sisi miring sama dengan jumlah dari luas daerah masing-masing kaki. Pembuktian teorema Pythagoras yang lainnya, terdapat pada penelitian [3] yang menjelaskan pembuktian teorema berdasarkan Geometri Euclid.

Pembuktian teorema kemudian dapat dikembangkan dalam konstruksi Pythagoras, seperti pada [4] yang menkonstruksi Pythagoras berdasarkan *Primitive Triple Pythagoras*. Selain itu, konstruksi Pythagoras dapat divisualisasikan ke dalam bentuk Pohon Fraktal seperti dalam artikel yang dibuat [5], dimana berdasarkan bukti teorema mengenai luas area persegi dari masing-masing sisi dapat membentuk pohon Pythagoras.

Konsep pohon Pythagoras inilah yang menjadi alternatif lain sebagai pengganti pohon hirarki dalam konstruksi pohon keputusan [6], dimana dengan adanya pohon fraktal ini dapat menghemat tempat dibandingkan menggunakan pohon hirarki. Hal ini menjadi dasar penelitian [7] tentang bagaimana membangun Pythagoras tersebut sebagai visualisasi data dari pohon keputusan di dalam Data Mining.

Pembuatan pohon keputusan di dalam Data Mining itu sendiri, memerlukan metode-metode algoritma yang di antaranya ID3, C4.5, J48, CART dan C5.0. Contoh eksperimental yang sering dijumpai seperti tulisan dari [8] menerapkan C4.5 sebagai pohon keputusan, dimana *Accuracy* sebanyak 98,80%, *Precision* 98,02%, dan *Sensitivity* 99,00% untuk memprediksi siswa dalam memperoleh bantuan dana pendidikan. Algoritma yang sama juga dilakukan pada penelitian [9] menggunakan C4.5 untuk kasus tanah longsor di daerah Purwakarta. Selain C4.5, penelitian dari [10] menggunakan metode J48 untuk menentukan besarnya area gempa bumi di Sumatera Utara.

Penelitian-penelitian yang telah disebutkan, menunjukkan bahwa metode algoritma tersebut membentuk pohon keputusan untuk menentukan aturan yang mempunyai hubungan asosiasi dengan pernyataan IF-THEN, sehingga data-data yang dikumpulkan dapat diklasifikasikan. Karena itu, ID3, C4.5, J48, dan C5.0 dapat dikatakan algoritma klasifikasi karena data-data masih dapat dikategorikan. Namun, beda halnya data-data yang bersifat kontinu, sehingga diperlukan algoritma regresi [11] seperti pada algoritma CART (walaupun masih dapat dikelompokkan sebagai algoritma klasifikasi). Konsep algoritma klasifikasi itu sendiri menggunakan *Gain* untuk mendapatkan nilai *Entropy* sehingga dapat diukur seberapa baiknya dari node pohon keputusan. Oleh karena itu, pada penelitian [12] menggunakan *Standard Deviation Reduction* dalam menentukan nilai node terbaik untuk membangun pohon keputusan dengan memasukkan konsep algoritma ID3.

Konstruksi Pythagoras Tree diterapkan oleh penulis untuk menentukan keputusan pada masalah kebiasaan penggunaan Instagram [13]. Data yang digunakan sebagai dasar konstruksi Pythagoras adalah *Initials*, *Age*, *Propose to use Instagram*, *Download Content*, *Frequency of using Instagram*, *The duration of using Instagram*, dan *Cumulative of Using Instagram*. Algoritma yang digunakan adalah *Standard Deviation Reduction* ID3, untuk mengklasifikasi data yang sebagian besar dapat dikategorikan dengan nilai target yang bersifat kontinu maupun numerik. Konstruksi Pythagoras tersebut menghasilkan visualisasi pohon keputusan berdasarkan ujung batang pohon yang disebut dengan daun. Aturan dari pohon Pythagoras tersebut didapatkan bahwa terdapat nilai target yang dapat diklasifikasikan yaitu *Age* 19, namun terdapat nilai atribut yang masih ambigu dikarenakan data yang sedikit, sehingga aturan yang didapatkan menjadi kurang jelas.

Oleh karena itu, pada kasus ini penulis membangun Pythagoras dalam memvisualisasikan *Random Forest* untuk pemodelan pohon keputusan dengan tujuan dapat memudahkan penelusuran atribut yang tadinya pada pohon keputusan dari [13] terdapat data ambigu. [14] menggunakan *Random Forest* pada penelitiannya untuk mengatasi data yang tidak seimbang untuk studi kasus Penilaian Mutu Pendidikan. Hal ini dapat terjadi karena ada kelas data yang sedikit sehingga terjadi kesalahan klasifikasi. Pada penelitian lain, yaitu [15] menyatakan bahwa *Random Forest* dapat digunakan untuk klasifikasi dan regresi, tujuan penelitian ini untuk meningkatkan akurasi terutama pada studi kasus curah hujan. *Random Forest* sebenarnya terdiri dari beberapa pohon keputusan yang dibundel menjadi satu. Setiap pohon keputusan yang dibuat, dilakukan pemisahan atribut secara acak dan juga dilakukan pemilihan acak pada *instance* untuk membangun pohon keputusan. Kemudian dilakukan observasi untuk dilakukan *voting* untuk memilih pohon keputusan yang memiliki keakuratan yang baik.

Adapun formula yang digunakan dalam membangun pohon keputusan, berdasarkan dari [7] dengan menggunakan Standar Deviasi ( $\sigma$ ), yaitu:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (u_i - \mu)^2} \quad (1)$$

dimana

$$\mu = \frac{1}{n} \sum_{i=1}^n u_i \quad (2)$$

$\sigma$  digunakan untuk membangun pohon, yaitu dengan memisahkan subset sesuai dengan atribut yang dipilih secara random sehingga membentuk cabang. Kemudian rata-rata ( $\mu$ ) sebagai nilai dalam daun (*leaf node*), biasanya nilai ini merupakan suatu nilai dari kelas target. Selain itu, untuk menghentikan pembentukan cabang menggunakan Koefisien Variasi (KV) berdasarkan ambang batas  $\leq 0,1$ . KV ini juga sebagai pembandingan cabang mana yang layak di gunakan untuk pembentukan pohon keputusan dalam bentuk prosentase. Berikut rumus KV:

$$KV = \frac{\sigma}{\mu} \times 100\% \quad (3)$$

Seperti penjelasan sebelumnya, pohon keputusan dapat divisualisasikan dengan menggunakan Pythagoras dimana rumusnya adalah

$$a^2 + b^2 = c^2 \quad (4)$$

Jika  $c^2$  adalah subset dari *root node*,  $a^2$  dan  $b^2$  adalah *leaf node*, maka nilai sisi miring adalah akar kuadrat dari *root node* dan nilai kaki adalah akar kuadrat dari *leaf node*. Komponen  $a^2$ ,  $b^2$  dan  $c^2$  juga dapat dianggap sebagai luas bujur sangkar sehingga masing-masing sisi dapat dilihat pada persamaan (5):

$$L = a^2, b^2, c^2$$

$$\sqrt{L} = a, b, c \tag{5}$$

**2. METODE**

Penulis membangun Pythagoras dalam visualisasi *Random Forest* untuk menelusuri atribut yang telah dibangun menjadi pohon keputusan dengan bantuan *Software* yaitu *Orange Data Mining*. Tahap-tahap yang dilakukan dalam penelitian ini adalah sebagai berikut:



Gambar 1. Tahap-tahap Pembangunan Pythagoras dalam Penelusuran Atribut Bermasalah

Berdasarkan pada Gambar 1, sebelum membangun Pythagoras adalah memilih *attribute* dan *instance* secara acak. Sebagai contoh dapat dilihat pada Gambar 2, yang dimana terdapat 4 dari 5 atribut yang dipilih yaitu *Propose to use Instagram*, *Download Content*, *Frequency of using Instagram*, dan *Cumulative of Using Instagram* untuk kelas target *Age* (Usia). Begitu juga dengan *Instance* yang dipilih secara acak, contohnya nomer 2, 3, 5, 7, 9, dan 13. Oleh karena itu, 33 instance yang merupakan total keseluruhan data dalam penelitian ini, tidak semua akan diambil untuk dibangun menjadi pohon keputusan.

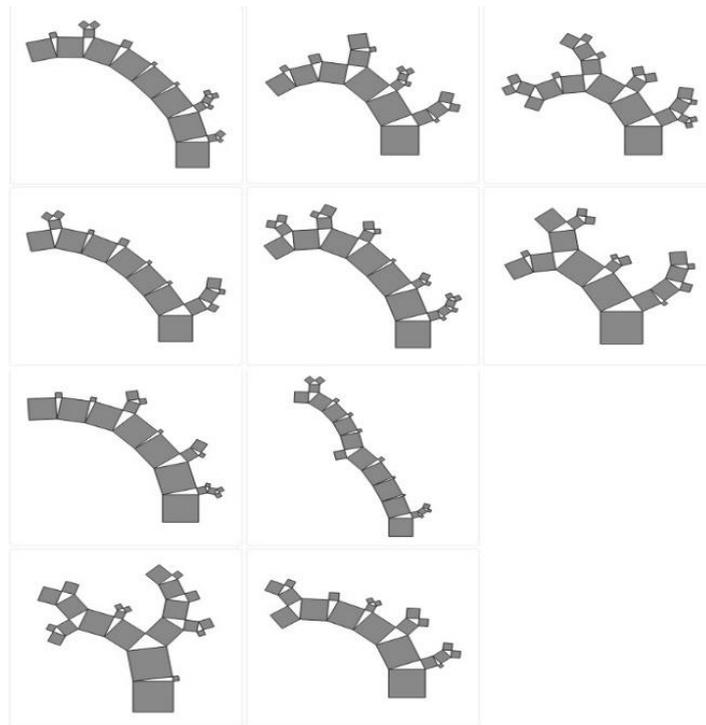
	Initials	Age	Propose to use Inst.	Download Conten	Frequency of using Inst.	Duration of using Ins	Cumulative of Using Inst
1	A1	20	Communicate l...	Another person...	1 to 3 times	< 5 minutes	<0.5 hour
2	A2	21	Uploading edit...	Video cinematic	1 to 3 times	< 5 minutes	<0.5 hour
3	A3	19	Communicate l...	Group Photo	1 to 3 times	5 to 15 minutes	0.5 to 1 hour
4	A4	?	Uploading Con...	Selfie	1 to 3 times	30 to 60 minutes	1 to 2 hours
5	A5	19	Uploading Con...	Selfie	1 to 3 times	30 to 60 minutes	1 to 2 hours
6	A6	18	Communicate l...	Selfie	>9 times	30 to 60 minutes	1 to 2 hours
7	A7	19	Uploading Con...	Objects Photo	4 to 6 times	5 to 15 minutes	0.5 to 1 hour
8	A8	19	Communicate l...	Selfie	>9 times	15 to 30 minutes	1 to 2 hours
9	A9	18	Posting and Inf...	Objects Photo	4 to 6 times	5 to 15 minutes	2 to 4 hours
10	A10	19	Surfing	Objects Photo	1 to 3 times	15 to 30 minutes	2 to 4 hours
11	A11	21	Surfing	Selfie	1 to 3 times	< 5 minutes	<0.5 hour
12	A12	19	Communicate l...	Selfie	7 to 9 times	30 to 60 minutes	1 to 2 hours
13	A13	19	Surfing	Selfie	1 to 3 times	15 to 30 minutes	0.5 to 1 hour
14	A14	18	Uploading Con...	Selfie	?	>60 minutes	4 to 6 hours
15	A15	18	Surfing	Selfie	>9 times	5 to 15 minutes	4 to 6 hours
16	A16	19	Posting and Inf...	Objects Photo	4 to 6 times	15 to 30 minutes	1 to 2 hours

Gambar 2. Contoh Pemilihan atribut dan *Instance* secara acak

Kemudian, membangun Pythagoras menjadi pohon keputusan berdasarkan jumlah sampel berdasarkan *instance*. Pada penelitian ini, hanya membuat 10 pohon keputusan yang kemudian diselusuri untuk menemukan atribut bermasalah. Maksud atribut yang bermasalah ini adalah atribut ambigu yang ditemukan pada penelitian sebelumnya [13]. Atribut ini dijadikan kelas target sebagai eksperimen, yang kemudian dibandingkan dengan kelas target sebelumnya (kelas target *Age* atau usia). Setelah itu, baru dilakukan analisa untuk mengetahui penyebab kenapa atribut tersebut ada masalah.

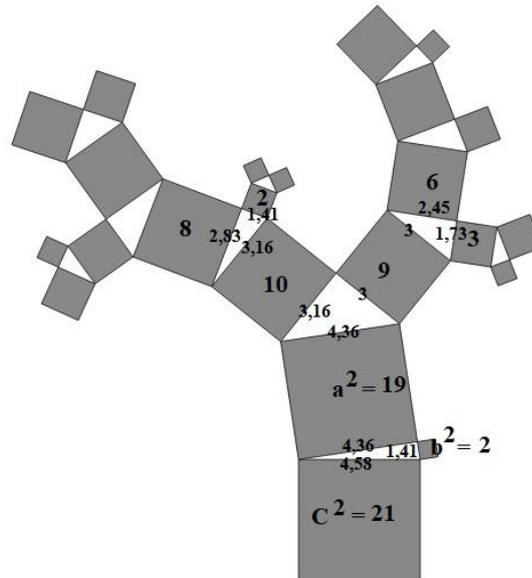
### 3. HASIL DAN PEMBAHASAN

Seperti dijelaskan pada bagian Metode, penelitian ini mengambil 33 instance dan 5 atribut dengan melakukan pemilihan data secara acak, sehingga terbentuk 10 pohon yang diambil dari penelitian yang pernah penulis buat yaitu masalah kebiasaan penggunaan Instagram. Berikut adalah pohon-pohon keputusan yang dibuat dalam *Random Forest*:



Gambar 3. Pohon-pohon dalam *Random Forest*

Pohon-pohon pada Gambar 3, dibangun berdasarkan rumus (4) dimana  $c^2$  sebagai *root node* (garis miring) dan  $a^2$ ,  $b^2$  sebagai *leaf node* (kaki). Sebagai contoh: jika  $c^2 = 21$ , kemudian dilakukan pemisahan data dengan menggunakan  $\sigma$  sehingga menghasilkan  $a^2 = 19$  dan  $b^2 = 2$ . Maka nilai sisi berdasarkan rumusan (5) dari masing-masing  $a$ ,  $b$ ,  $c$  adalah 4,36; 1,41; 4,58. Pembentukan Pythagoras dapat digambarkan dari salah satu pohon di bawah ini:



Gambar 4. Pythagoras Pada Salah Satu Pohon *Random Forest*

Pembangunan Pythagoras pada Gambar 4, berdasarkan dari jumlah *instance* secara acak. Nilai  $c^2$  sebagai awal *root node* yang mempunyai 21 *instance*. Kemudian dibagi menjadi 2 kaki ( $a^2$  dan  $b^2$ ) yaitu 19 *instance* dan 2 *instance*. Jika dijumlahkan kedua kaki tersebut menjadi 21 *instance*. Begitu juga dengan nilai-nilai lain, menjadi  $(a^2, b^2, c^2) = (10,9,19)$ ,  $(8,2,10)$ ,  $(6,3,9)$ , dan seterusnya; sampai terbentuknya pohon keputusan. Pembangunan Pythagoras yang dibuat dalam pohon *Random Forest*, memudahkan penelusuran data dalam atribut jika dibandingkan dengan [13], seperti pada gambar di bawah ini:



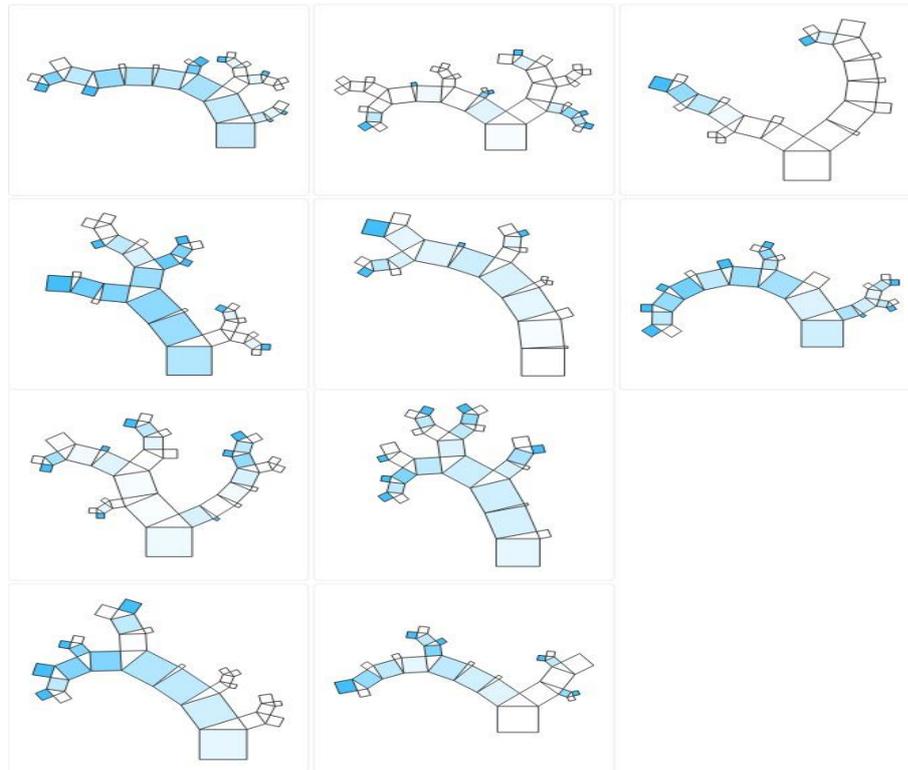
Gambar 5. Pemisahan Data Sampel Pada Cabang Pythagoras

Menurut gambar 5, untuk usia 20 terdapat tiga kali pemisahan data pada atribut yang sama, yaitu *Cumulative of Using Instagram* (kumulatif penggunaan Instagram). Pemisahan data tersebut, antara lain:

1. 4 sampai dengan 6 jam dengan  $\mu=\text{nan}$  dan  $\sigma=\text{nan}$
2. 0,5 sampai dengan 1 jam dengan  $\mu=19,5$  dan  $\sigma=0,5$
3. 1 sampai dengan 2 jam dengan  $\mu=19,75$  dan  $\sigma=0,433$

Berdasarkan dari tiga kali pemisahan data, sehingga membuat klasifikasi menjadi ambigu, alasannya adalah karena terlalu banyak kriteria di dalam satu atribut. Hal ini mungkin terjadi jika tidak ditemukan hubungan antara subset data yang spesifik dalam atribut terhadap kelas target.

Oleh karena itu, atribut kumulatif penggunaan Instagram dibuat menjadi kelas target. Sebagai contoh eksperimen yang dilakukan, yang mana dengan tujuan menghasilkan pohon-pohon *Random Forest* agar dapat dianalisa mendalam dengan langkah-langkah yang sama dengan yang dilakukan pada tahapan pertama dalam pemilihan atribut dan *instance* secara acak. Hasil pembuatan pohon-pohon keputusan yang dibangun Pythagoras dapat dilihat sebagai berikut:



Gambar 6 *Random Forest* Kelas Target Kumulatif Penggunaan Instagram

Berdasarkan Gambar 6, ditemukan masih terdapat beberapa kriteria yang tidak mempunyai hubungan terhadap kelas target. Hal ini, dapat dilihat juga dari pewarnaan pada pohon yang dibangun oleh Pythagoras. Jika dibandingkan dari pewarnaan pada [7], seharusnya pewarnaan gradasi dari *root node* ke *leaf node*, dimulai dengan warna gelap ke terang yang menunjukkan adanya hubungan atribut dari masing-masing *instance*. Namun pada kasus kelas target Kumulatif Penggunaan Instagram, pewarnaan hubungan menjadi tidak beraturan karena teridentifikasi hanya mempunyai jenis warna biru dan putih. Tidak beraturan berarti tidak adanya hubungan subset data dari setiap atribut dalam *instance*. Karenanya, walaupun sudah membentuk Pohon Keputusan berdasarkan Pythagoras, pohon tersebut tidak dapat membuat *rule* sebagai hubungan antar atribut terhadap kelas target yang dapat mengakibatkan terjadinya kesalahan klasifikasi. Alasan lain terjadinya kesalahan klasifikasi dapat disebabkan karena pohon keputusan tersebut tidak mengenali pola data.

Kesalahan klasifikasi yang telah divisualisasikan dalam Pythagoras untuk *Random Forest* dapat dilihat pada *Confusion Matrix* di bawah ini sebagai informasi kesalahan prediksi, dimana data yang benar adalah 0.5 ke 1 jam sebanyak 27,3%, 1 sampai 2 jam sebanyak 37,5%, 2 sampai 4 jam 0%, 4 sampai 6 jam sebanyak 0%, dan <0.5 jam sebanyak 70% :

		Predicted					
		0.5 to 1 hour	1 to 2 hours	2 to 4 hours	4 to 6 hours	<0.5 hour	$\Sigma$
Actual	0.5 to 1 hour	27.3 %	50.0 %	50.0 %	50.0 %	20.0 %	11
	1 to 2 hours	27.3 %	37.5 %	50.0 %	0.0 %	0.0 %	7
	2 to 4 hours	18.2 %	12.5 %	0.0 %	50.0 %	0.0 %	4
	4 to 6 hours	9.1 %	0.0 %	0.0 %	0.0 %	10.0 %	2
	<0.5 hour	18.2 %	0.0 %	0.0 %	0.0 %	70.0 %	9
$\Sigma$		11	8	2	2	10	33

Gambar 7 *Confusion Matrix* Untuk Kelas Target Kumulatif Penggunaan Instagram

Dengan kata lain hanya data <0.5 yang mendekati ketepatan prediksi data. Kesalahan klasifikasi pada penelitian ini, data ambigu ternyata diakibatkan karena data tersebut tidak valid, bukan hanya karena data yang sedikit. Namun berdasarkan dari pembangunan Pythagoras untuk Random Forest, data-data tersebut lebih mudah ditelusuri kesalahannya, yang mana tidak dapat dilakukan oleh pohon keputusan tunggal.

#### 4. PENUTUP

Berdasarkan analisa dari Hasil dan Pembahasan, dapat disimpulkan data-data yang dibangun dengan menggunakan Pythagoras sebagai konstruksi pohon keputusan, yang sudah dilakukan pada penelitian mengenai kebiasaan penggunaan Instagram yang pernah dilakukan penulis sebelumnya. Penggunaan *Random Forest* menyebabkan data tersebut dapat ditelusuri berdasarkan pembangunan Pythagoras dengan cara mengambil *instance* secara acak. Eksperimen dilakukan dengan menggunakan Usia sebagai kelas target, sehingga hasil yang didapatkan adalah 3 kali pemisahan data pada atribut kumulatif penggunaan Instagram, sebagai contoh pada eksperimen ini adalah usia 20. Pemisahan data dari kumulatif penggunaan Instagram tersebut antara lain didapatkan 4 sampai dengan 6 jam dengan nilai  $\mu=\text{nan}$  dan  $\sigma=\text{nan}$ ; 0,5 sampai dengan 1 jam dengan  $\mu=19,5$  dan  $\sigma=0,5$ ; 1 sampai dengan 2 jam dengan  $\mu=19,75$  dan  $\sigma=0,433$ . Pemisahan tersebut, jika dilihat dari data yang dibangun didalam Pythagoras, teridentifikasi bahwa data tersebut menjadi ambigu yang disebabkan beberapa subset tidak punya hubungan, sehingga data menjadi tidak valid atau dengan kata lain tidak dapat mengenali pola. Kemudian, Penulis juga merubah kelas target yaitu kumulatif penggunaan Instagram dan menemukan tidak ada hubungan antar atribut berdasarkan pewarnaan pada pohon. Kemudian, berdasarkan pada *Confusion Matrix* bahwa keakuratan prediksi data yang didapatkan sangatlah sedikit, antara lain 0.5 ke 1 jam sebanyak 27,3%, 1 sampai 2 jam sebanyak 37,5%, 2 sampai 4 jam 0%, 4 sampai 6 jam sebanyak 0%, dan <0.5 jam sebanyak 70%. Oleh karena itu, kesimpulan yang didapatkan dari *Confusion Matrix*, ternyata memang terjadi kesalahan klasifikasi. Namun, data-data tersebut mudah ditelusuri lebih dalam dimana letak kesalahannya dengan membangun Pythagoras pada *Random Forest*, yang mana tidak dapat dilakukan oleh pohon keputusan tunggal.

Berdasarkan pembangunan Pythagoras yang penulis buat, diharapkan untuk penelitian selanjutnya dapat dilakukan observasi dengan menggunakan *Gain* untuk mendapatkan nilai *Entropy* dalam membangun Pythagoras pada Pohon Keputusan berdasarkan *Random Forest*. Selain itu juga dapat membandingkan antara Pythagoras Regresi dengan Pythagoras Klasifikasi, dalam pembentukan pohon keputusan.

#### DAFTAR PUSTAKA

- [1] B. Ratner, "Pythagoras: Everyone knows his famous theorem, but not who discovered it 1000 years before him," *J. Targeting, Meas. Anal. Mark.*, vol. 17, no. 3, pp. 229–242, Sep. 2009.
- [2] S. Zaerani, M. Mardhiah, and S. Suharti, "Pengaruh Penguasaan Konsep Teorema Pythagoras Terhadap Kemampuan Menyelesaikan Soal-Soal Bangun Ruang Sisi Datar Pada Siswa Kelas Viii Mts Negeri Balang-Balang," *MaPan*, vol. 5, no. 2, pp. 279–292, 2017.
- [3] M. Naufal Faris, S. Ulfa, and H. Praherdhiono, "TEKNOLOGI PEMBELAJARAN MATEMATIKA PEMBUKTIAN TEOREMA PYTHAGORAS BERBASIS VISUAL," *JINOTEP (Jurnal Inov. dan Teknol. Pembelajaran) Kaji. dan Ris. dalam Teknol. Pembelajaran*, vol. 6, no. 1, pp. 8–14, 2019.
- [4] M. Affaf, "Konstruksi baru untuk tripel pythagoras," *J. Pendidik. dan Pembelajaran Mat.*, vol. 2, no. 1, pp. 69–78, 2016.
- [5] J. N. da Costa Reis, "Uma Árvore de Pitágoras Explorando os Fractais no Ensino Médio," *Ciência e*

- Nat.*, vol. 37, no. 3, pp. 411–418, 2015.
- [6] F. Beck, M. Burch, T. Munz, L. Di Silvestro, and D. Weiskopf, “Generalized pythagoras trees: A fractal approach to hierarchy visualization,” in *Communications in Computer and Information Science*, 2015, vol. 550, pp. 115–135.
- [7] E. W. Ambarsari, A. Ar Rakhman Awaludin, A. Suryana, P. M. Hartuti, and R. Rahim, “BASIC CONCEPT PYTHAGORAS TREE FOR CONSTRUCT DATA VISUALIZATION ON DECISION TREE LEARNING,” *J. Appl. Eng. Sci.*, vol. 17, no. 4, pp. 468–472, 2019.
- [8] S. Hendrian, “ALGORITMA KLASIFIKASI DATA MINING UNTUK MEMPREDIKSI SISWA DALAM MEMPEROLEH BANTUAN DANA PENDIDIKAN,” *Fakt. Exacta*, vol. 11, no. 3, pp. 266–274, 2018.
- [9] Y. Handrianto and M. Farhan, “C.45 Algorithm for Classification of Causes of Landslides,” *SinkrOn*, vol. 4, no. 1, pp. 120–127, Oct. 2019.
- [10] D. Parwatingtyas and E. W. Ambarsari, “CALCULATION ANALYSIS TO DETERMINE THE LARGEST EARTHQUAKE AREA AT NORTH SUMATERA USING J48 DECISION TREE METHOD,” in *The Proceeding of Ocean, Mechanical and Aerospace -Science and Engineering-(POMase)*, 2016, vol. 3, no. 1, pp. 36–43.
- [11] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification And Regression Trees*. Routledge, 2017.
- [12] E. Windia Ambarsari, R. Avrizal, E. Doni Sirait, S. Dwiasnati, and R. Rahim, “Regression Tree Role for Interpret Monetizing of Game Live Streaming,” *J. Phys. Conf. Ser.*, vol. 1424, p. 012014, Dec. 2019.
- [13] E. W. Ambarsari and H. Herlinda, “Pythagoras Tree Applied For Determined Instagram Usage Habit Decision,” *SinkrOn*, vol. 4, no. 2, pp. 56–61, 2020.
- [14] A. Ramadhan, B. Susetyo, and I. Indahwati, “PENERAPAN METODE KLASIFIKASI RANDOM FOREST DALAM MENGIDENTIFIKASI FAKTOR PENTING PENILAIAN MUTU PENDIDIKAN,” *J. Pendidik. dan Kebud.*, vol. 4, no. 2, pp. 169–182, Dec. 2019.
- [15] A. Primajaya and B. N. Sari, “Random Forest Algorithm for Prediction of Precipitation,” *Indones. J. Artif. Intell. Data Min.*, vol. 1, no. 1, pp. 27–31, Mar. 2018.