Application of the C4.5 Algorithm on the Effect of Watching Youtube Videos On the Development of Early Childhood Creativity

Ita Dewi Sintawati[1)]*, **Widiarina[2)]**, **Kartika Mariskhana[3)]**
[1)2)3)] Universitas Bina Sarana Informatika, Jakarta, Indonesia
[1)] ita.ids@bsi.ac.id, [2)] widiarina.wda@bsi.ac.id, [3)]kartika.kma@bsi.ac.id

Abstract: Youtube media is one of the social media for communication used by the community, not out of reach of children. The rapid development of information technology, of course, this has an influence on human life. Talking about life in humans, it cannot be separated from human behavior. Some psychologists say that children tend to fully absorb what they see, and children will learn from what they see. This can trigger creativity for young children. The creativity of each individual can be seen in terms of how he makes something he thinks of because he sees an object that already exists and then he innovates it into a new form. This attracted the attention of the author to identify and describe the impact of watching YouTube videos on the development of early childhood creativity. This type of research is based on developing phenomena, how much influence is brought about by technological advances on YouTube social media in the formation of children's behavior. The process of completing the goals to be achieved in this study is to provide information about recommendations for child development with positive creativity, making it easier to determine early childhood development by using the Decision Tree Algorithm C4,5 method. The problem in this study is that early childhood imaginations are higher and will be affected by streaming video ads on YouTube. The results showed that children aged 3 and 4 years often watched, while children aged 5 and 6 years did not watch often, so the role of parents was more dominant in supervising children aged 3 and 4 years.

Keywords: Early Childhood, Decision Tree, Youtube C4.5 Algorithm,Entropy, Gian

INTRODUCTION

With the development of today's technology, anything and anyone who is far away will be closer if connected to the internett(Mastanora, 2018). The internet is known to be superior in gathering various people, with any geographical conditions it will not be a limiting problem (Nikmah, 2017). Video-based social media that is often accessed on smartphones is Youtube. Starting from news, Music with the latest clips, Comedy Videos or Animations, everything is on Youtube. Youtube is intended for those who need audio and visual information, visitors to this site can also upload their videos to Youtube and share them all over the world(Irawan, 2018). Youtube is part of the New Media which is accessed using the Internet is a platform for publishing videos(Irawan, 2018). This platform can be used by anyone and anywhere as long as it has an internet network(Triandini, 2019). Youtube also provides a forum for people to connect and share information and inspiration around the world(David, 2017). Youtube is the leader for video search sites on the internet, with more than 100,000,000 videos being watched by visitors every day. More than 65,000 videos are now uploaded daily to Youtube(Putra, 2018). Most people perceive that YouTube is an application that requires everyone to have an account in order to be free to access the application(Nur, 2019), this has become a basic need for everyone of all ages, from parents who already have grandchildren to even young children. become addicted to watching YouTube videos, some are downloaded online or offline, some are watching movies live streaming. The impact of watching this youtube

*name of corresponding author

video is of course various, if you see a video with positive content it will have a positive effect, but if you see a video with negative content it will cause negative things for the audience, it can even lead to criminal acts if misused , because anyone can share videos with any content. For this reason, early childhood children need special escort if they are going to be given watching videos on the YouTube application. Humans socially in mass communication tend to learn about what is a stimulus from their environment(Mastanora, 2018), for example, after seeing it in audio-visual form. In the learning process through observation there are several elements of learning, namely attention, recall, repetition, and motivation. children learn and imitate the behavior they observe. In the science of child psychology, some experts say that children tend to fully absorb what they see, and they will learn from what they see. This can trigger creativity for young children. The creativity of each individual can be seen in terms of how he makes something he can think of because he sees an object that already exists and then he innovates it into a new form(Khasanah, 2019). This attracted the author's attention to identify and describe the impact of watching YouTube videos on early childhood(Gita, 2019).

The hypothesis is a temporary answer to the problem that is the object of research. In this study, the author discusses the C4.5 Algorithm Analysis in determining the factors that influence the emergence of video advertisements in early childhood development. The types of hypotheses in the study can be described as follows::

1. Has no influence on early childhood development on streaming video ads on YouTube using the C4.5 Algorithm method.
2. Having an influence on early childhood development on streaming video ads on YouTube using the C4.5 Algorithm method.

LITERATURE REVIEW

Several previous studies with the application of the Decision Tree method are as follows:

1. The purpose of the study was to determine the effect of youtube-based e-learning media on student learning outcomes in real analysis courses, this study used a quasi-experimental method with research design and sampling using purposive sampling technique, where the sample was divided into 2 classes, namely the experimental class and the first class. control. The results showed that the average student learning outcomes using youtube-based e-learning media were higher than those without using youtube-based e-learning media, namely 65.9 for the experimental class and 50.4 for the experimental class. (Panggabean, 2018).
2. The research method used is descriptive research method where the data will be processed to make a systematic, factual, and accurate description of the facts or characteristics of a particular population. This study aims to examine the level of popularity of several youtube channels from the number of uploads and subscribers and also the number of people who view the video. The results show that the level of popularity (which is judged based on views, subscribers, and uploads) of a total of 5001 different youtube channels can be in the following order: A++,A+,A,A-,B+(Sudarsono, 2020).
3. Research on Youtube channel ratings can be judged by Video Uploads, Subscribers, Video Viewers. From Youtube data processing can be done. One of the data processing techniques that can be used in the process is classification. Classification is a data processing technique that divides objects into classes according to the desired number of classes. And using the C4.5 algorithm in the classification process. Who can determine the Youtube Channel, especially Indonesia with a Very Good or Good ratio(Wahyudi & Dkk, 2019).

METHOD

The research carried out is a step to achieve the objectives of the results, the stages of the research are as follows::

1. Data collection is a very important stage in a study, because the data collected is the most important part of an object of research. At the stage of data collection, the authors collect the data needed to be used as research material, including observations, interviews and literature(Shidiq, 2019). Population is a generation area consisting of objects or subjects that have certain qualities and characteristics that are applied by researchers to be studied and then drawn conclusions. The population of this study is the response of parents to the influence of early childhood development who use YouTube a lot nowadays where there are video streaming advertisements on YouTube(GULO, 2020). The sample is part of the number and characteristics possessed by the population. Based on the sample taken in the population that has been described, it is determined that the number of parents' responses is 100 data on the influence of early childhood development on the emergence of streaming video ads on YouTube.
2. The stages of analysis in this study were carried out with the data needed to be processed into important data, where the research went through the following stages:
   a. Data Needs Analysis

*name of corresponding author

This stage is the stage in analyzing the data that has been collected and will be used in research. The analysis of data needs for this study is to collect data on families who have early childhood to find out how much influence the child's development has on the emergence of streaming video ads on YouTube that the child watches(Winarni, 2017). Collecting data as many as 100 early childhood children in the area where the research writer is located.

b. Process Method Algorithm C4.5

After the author determines the data requirements, then the next document will be processed through the processing stage with the C4.5 Algorithm method. At this stage classify objects based on training samples, training labels, attributes(Tambunan, 2020). The calculation is to find the gain first in order to get the root node we want. The following is the flow at the stages of the process using the C4.5 Algorithm method. The technique of making a decision tree, Research looks for attributes first to determine the root of the decision tree, then calculates the total entropy, entropy of the attributes of each value and then calculates the gaint of the entropy of the attribute to make it easier to determine the problem that the researcher makes.

3. Testing Stage

At this stage of research, the authors conducted testing using the following stages:

a. Excel Test

Testing at this stage the author performs data processing using Microsoft Excel and calculations manually.

b. Rapidminer Test

At this stage the author uses the Rapid Miner application software, in the software the calculation is available automatically.

Table 1. Early Childhood Data

Number	Parents' name	Child Age	Comprehension	Watching Often	The role of parents
1	Astri Yani	4	Intermediate	Not often	Yes
2	Apriyanti	3	Intermediate	Not often	Yes
3	Jayadih	3	Intermediate	Not often	Yes
4	Ari Dewi	4	Intermediate	often	Yes
5	Sari Avril	6	Good	Often	No
6	Yuli	6	Good	Often	No
7	Hendro	4	Intermediate	often	Yes
8	Fenny	5	Good	Often	No
9	Sarah	3	Intermediate	Not often	Yes
10	Neneng	5	Good	often	No
Up to data to 100					
100	Santi	3	Medium	Not often	Yes

*name of corresponding author

RESULT

In the development of the Data Mining Algorithm C4.5 method, it is based on research. Where in this study the influence of early childhood development on streaming video ads on youtube which is assessed based on the criteria to be tested using the C4.5 algorithm method. The calculation process looks for the root of the problem by determining the highest Entropy and Gain from each calculated result and determining the root.

Table 2. Roots of Decision

Node	Attribut	Value	Number of Cases	NO(S1)	YES(S2)	Entropy	Gain
1	total		100	42	58	0,9814539	
	Child Age						0,592175
		3	23	0	23	0	
		4	24	0	24	0	
		5	26	20	6	0,7793498	
		6	27	22	5	0,6912899	
	Comprehension						0,054999
		Intermediate	47	13	34	0,8507707	
		Good	53	29	24	0,9935705	
	Watching Often						0,254191
		Often	53	37	16	0,8835851	
		Not often	47	6	41	0,5509844	

In table 2. From several variables, the child's age variable becomes the root of the tree with an Entropy value of 0.9814539 and a Gian value of 0.592175.

Table 3. Node Calculation 1.1

Node	Attribut	Value	Number of Cases	NO(S1)	YES(S2)	Entropy	Gain
1	total		100	42	58	0,9814539	
	Child's Age = 6 years						0,5921747
		6	23	0	23	0	
		7	24	0	24	0	
		8	26	20	6	0,7793498	
		9	27	22	5	0,6912899	
	Comprehension						0,0549993
		Intermediate	47	13	34	0,8507707	
		Good	53	29	24	0,9935705	
	Watching Often						0,2541911
		Often	53	37	16	0,8835851	
		Not often	47	6	41	0,5509844	

*name of corresponding author

In table 3, look for Node 1.1. after making the results of the calculation of the table data above, the highest gain is in the attribute of Frequent Watching at the age of 6 years old.

Table 4. Calculation of Node 2

Node	Atribut	Value	Number of Cases	No (Si)	Yes (Si)	Entropy	Gain
1.2	Child Age = 5 Years		26	20	6	0,7793498	
							0,1218448
	Comprehension						
		Intermediate	7	7	0	0	
		Good	19	13	6	0,8997438	
							0,7793498
	Watching Often						
		Often	20	20	0	0	
		Not often	6	0	6	0	

In Table 4, based on the results of the above data calculation, the highest gain is in the frequency of watching (Children's Age Attribute = 5). Then the branch of the decision tree of 1.2 is Frequently watched.

The process of using excel calculations has determined the results, then for further processing the data with Rapidminer software to get the decision tree results, as evidence that what has been calculated through excel is correct and the results obtained are in accordance with the running process on rapidminer.

Figure 1. Excel data connection with Rapidminer

Figure 1. The next process is connecting excel data by forming a decision tree with data in the form of variables that have been prepared in excel data by processing with rapidminer, the results of the decision tree will be shown in Figure 2

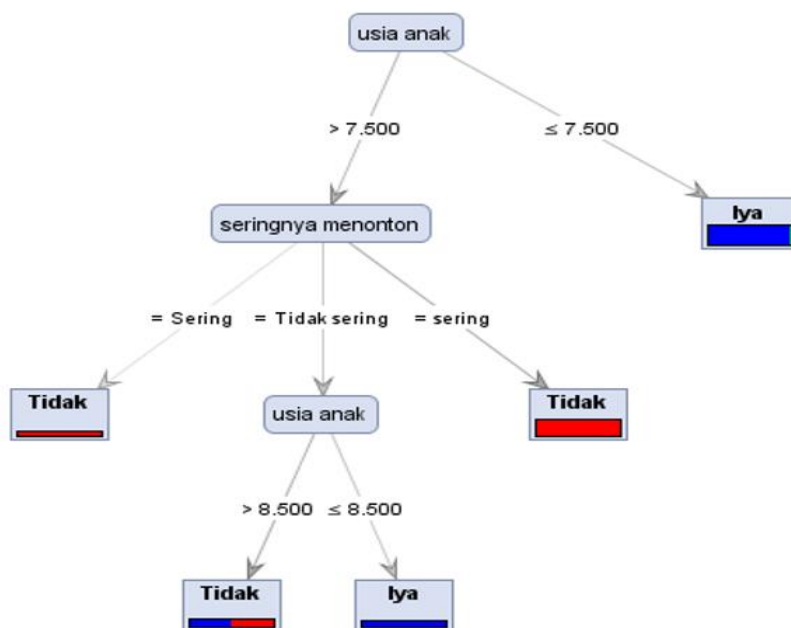*name of corresponding author

Figure 2. Decision tree with Rapidminer

Description in Figure 4. The formation of a decision tree with the variable age of the child as the root of the tree and forming the leaf level for children aged 3 and 4 years. For children aged 4 and 5 years forming the next leaf level with the variable watching often.

DISCUSSIONS

Based on the calculation process using Excel and data processing with Rapidminer software, the variable Often watching becomes a root node with the formation of a leaf node in children aged 3 and 4 years, it can be said that children aged 3 and 4 years are more dominant in viewing YouTube videos compared to children aged 5 and 6 years, this is also seen in daily activities, that children aged 3 and 4 years do activities in the Play Group where children play, while for children aged 5 and 6 years the place of activity is in advanced kindergarten, where they not only play but have been introduced to learning to read and count even with the media of pictures. From the results of research, early childhood can be directed to YouTube videos that can add creativity that can spur brain development, so that the positive things they access videos under the direction of parents, then brain intelligence for early childhood can develop well.

CONCLUSION

Based on this research, the conclusions obtained are that research on the influence of early childhood development on video streaming ads on YouTube can be applied using the calculation of the C4.5 Algorithm method. and rapidminer software data processing with detailed data analysis results, so that they can provide the right decisions. The role of parents in guiding early childhood to direct educative YouTube videos is more dominant in children aged 3 and 4 years, the role of parents who actively guide early childhood, it will be an advantage in itself in the development of the child's soul and become an individual. independent personality and creativity in positive activities. This writing aims to make parents aware of paying more attention to children at an early age, where the development of children at this age greatly affects the mindset of children in the future..

REFERENCES

David, E. R. (Eribka). (2017). Pengaruh Konten Vlog dalam Youtube terhadap Pembentukan Sikap Mahasiswa Ilmu Komunikasi Fakultas Ilmu Sosial dan Politik Universitas Sam Ratulangi. *Acta Diurna*, *6*(1), 93363. Retrieved from https://www.neliti.com/publications/93363/pengaruh-konten-vlog-dalam-youtube-terhadap-pembentukan-sikap-mahasiswa-ilmu-kom

Gita, L. (2019). KOMODIFIKASI SENSUALITAS DALAM TAYANGAN KIMI HIME DI MEDIA SOSIAL YOUTUBE. *SEMIOTIKA: Jurnal Komunikasi*, *13*(1), 89–105. https://doi.org/10.30813/s:jk.v13i1.1791

*name of corresponding author

GULO, Y. (2020). PENGARUH LINGKUNGAN KERJA, SELF-EFFICACY, DAN KARAKTERISTIK PESERTA PELATIHAN TERHADAP TRANSFER PELATIHAN KARYAWAN. *Media Bisnis*, *12*(1), 41–48. https://doi.org/10.34208/mb.v12i1.879

Irawan, R. E. (2018). Strategi NetTV Memproduksi Konten untuk Televisi Multiplatform. *Prosiding Konferensi Nasional Komunikasi*, *02*(01), 373–382.

Khasanah, Y. N. (2019). Meningkatan Kreativitas Melalui Kegiatan Kolase pada Anak Yuli. *Golden Age: Jurnal Ilmiah Tumbuh Kembang Anak Usia Dini*, *4*(1), 69–84. https://doi.org/10.14421/jga.2019.41-07

Mastanora, R. (2018). Dampak Tontonan Video Youtube Pada Perkembangan Kreativitas Anak Usia Dini Refika. *Jurnal Pendidikan Dan Perkembangan Anak*, *I*(2), 47–57.

Nikmah, F. (2017). KAJIAN TENTANG PEMASARAN ONLINE UNTUK MENINGKATKAN PELUANG BISNIS Oleh. *Adbis: Jurnal Administrasi Dan Bisnis*, *11*(1), 47. https://doi.org/10.33795/j-adbis.v11i1.19

Nur, M. I. (2019). Analisis Wacana Kritis pada Konten "Masjid untuk Semua" di Media Youtube CISFORM UIN Sunan Kalijaga, Yogyakarta. *Lentera*, *3*(1), 83–104. https://doi.org/10.21093/lentera.v3i1.1670

Panggabean, S. (2018). PENGARUH MEDIA E-LEARNING BERBASIS YOUTUBE TERHADAP HASIL BELAJAR MAHASISWA PADA MATA KULIAH ANALISIS REAL PRODI PENDIDIKAN MATEMATIKA FKIP UMSU. *Seminar Nasional Matematika Dan Terapan*, (1), 25–30. Retrieved from https://bulletin.indoms-acehsumut.org/index.php/simantap/article/view/144

Putra, A. (2018). PENGARUH YOUTUBE DI SMARTPHONE TERHADAP PERKEMBANGAN KOMUNIKASI INTERPERSONAL ANAK Asaas. *Jurnal Penelitian Komunikasi*, *21*(2), 159–172. https://doi.org/10.20422/jpk.v21i2.589

Shidiq, U. (2019). METODE PENELITIAN KUALITATIF DI BIDANG PENDIDIKAN. In *Journal of Chemical Information and Modeling* (Vol. 53). Retrieved from http://repository.iainponorogo.ac.id/484/1/METODE PENELITIAN KUALITATIF DI BIDANG PENDIDIKAN.pdf

Sudarsono, B. G. (2020). SEGMENTASI POPULARITAS AKUN YOUTUBE MENGGUNAKAN METODE ID3. *JBASE - Journal of Business and Audit Information Systems*, *3*(2), 32–38. https://doi.org/10.30813/jbase.v3i2.2269

Tambunan, M. G. (2020). Prediksi Kepribadian DISC Pada Twitter Menggunakan Metode Decision Tree C4.5 dengan Pembobotan TF-IDF dan TF-RF Maulina. *E-Proceeding of Engineering*, *7*(1), 2725–2738.

Triandini, E. (2019). Metode Systematic Literature Review untuk Identifikasi Platform dan Metode Pengembangan Sistem Informasi di Indonesia. *Indonesian Journal of Information Systems*, *1*(2), 63. https://doi.org/10.24002/ijis.v1i2.1916

Wahyudi, I., & Dkk. (2019). Klasifikasi Channel Youtube Indonesia Menggunakan Algoritma C4.5. *Teknik Komputer BSI*, *V*(1), 135–138. https://doi.org/10.31294/jtk.v4i2

Winarni, D. S. (2017). Analisis Kesulitan Guru Paud dalam Membelajarakan IPA pada Anak Usia Dini Dyah. *Edu Sains: Jurnal Pendidikan Sains & Matematika*, *5*(1), 12. https://doi.org/10.23971/eds.v5i1.578

*name of corresponding author