Utilization of Decision Trees in Intelligence Virtual Assistant

Putri Nurhaliza - 13520066[1]
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
[1]13520066@std.stei.itb.ac.id

***Abstract*— Intelligence virtual assistant is an IT product that capable to interact with humans. Virtual assistants have ability to give appropriate answers and perform certain actions by understanding what the user intends. A virtual assistant analyzes considerations in making decisions with database as preferences. Because there are many possibilities, virtual assistants need algorithms as decision support tools, one of which is decision tree. The decision tree classifies the keywords from the user that relate to the current condition and then determines what should it do to respond back the user.**

***Keywords*—Artificial Intelligence, Decision Tree, Machine Learning, Natural Language Processing, Virtual Assistant.**

I. Introduction

The times lead scientists to innovate developing technology that is closer to humans' behaviors, with a combination of artificial intelligence and machine learning. One of these innovations is an Intelligence Virtual Assistant (which in this paper will be referred to as virtual assistant). One of the most influential inventions was the chatbot ELIZA developed by MIT professor Joseph Weizenbaum in the 1960s as the first natural language processing computer program. Until now, many companies are competing to build virtual assistants to achieve their own goals. Whether it is intended to help people in their daily life like a real assistant, or just to improve customer service on their products. Some well-known virtual assistants are Google Assistant, Siri, Alexa, and Cortana.

People use virtual assistants because they can help solving problems in real-time which enhances human capability and productivity. Virtual assistant can be helpful when it is integrated with Internet of Things (IoT). Virtual assistants are required to be smart and capable to interact with humans. It should understand what human say either via speech or text. The smarter it is, the more it will comprehend humans. Understanding natural language is a significant leap in the development of smarter machines.

To build and develop a virtual assistant, it is necessary to combine a lot of methods which support the virtual assistant functionality. In addition to the utilization of Natural Language Processing (NLP), also needed an effective algorithm on the virtual assistant brain so that it could make the best decision. One of the decision support tools, which is commonly used is Decision Tree. To provide optimal results, decision trees require robust databases as consideration in making the best decisions. This paper will discuss how is the utilization of Decision Tree to determine proper response of virtual assistant to user's command.

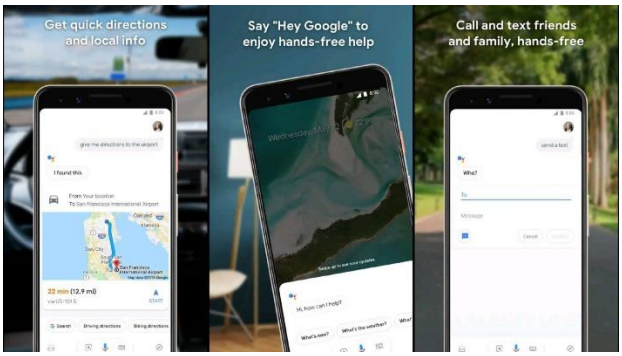


Fig. 1.1. Examples of Virtual Assistant: Google Assistant
(Source: https://play.google.com, accessed on 13 December 2021)

II. Basic Theory

A. Tree

A tree is an undirected connected graph that does not have a circuit in it. This means in a tree, every pair of nodes only have one path to connect them to each other. A forest is defined as a group of trees that are not connected to each other.

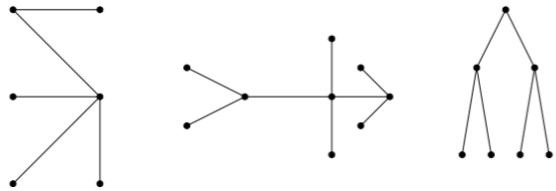
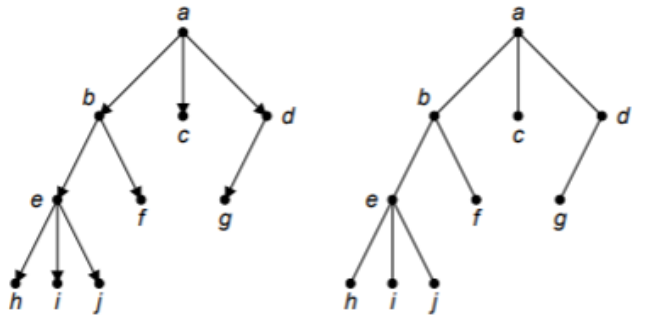


Fig. 2.1 Examples of Trees
(https://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2020-2021/Pohon-2020-Bag1.pdf, accessed on December 10, 2021)

A simple undirected graph G = (V, E) with the number of vertices is n is a tree if and only if [1]:

1. Each pair of vertices in G is connected by a single path
2. G is connected and has m = n – 1 edge
3. G contains no circuits and has m = n – 1 edge
4. G contains no circuit and adding one edge to the graph will create only one circuit
5. G is connected and all sides are bridges

There are some types of trees such as rooted tree, a tree where one of the nodes is treated a root and all the other edges in the tree are directed away from the root, turning the rooted tree into a directed graph, unlike a normal tree which is an undirected graph. As a convention, the arrow on the edge can be ignored. A rooted tree where all of its nodes have N or less children. called N-ary tree. The two special cases in the naming for N-ary tree are binary tree with the value of the N is equal to two and ternary tree with the value of the N is equal to three.



(a) (b)

Fig. 2.2. Example of rooted trees, (a) with arrow and (b) without arrow. (Source: https://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2020-2021/Pohon-2020-Bag2.pdf, accessed on December 10, 2021)

Tree has some important terms for each part of tree, they are: (the example below is corresponded to Fig. 2.2.)

1. Root: The top node of a tree (e.g., node ‘a’)
2. Parent: A predecessor one level above a node (e.g., node ‘b’ is parent of ‘f’
3. Child: A successor one level above a node (e.g., node ‘f’ is child of ‘b’)
4. Siblings: group of nodes that have same parent (e.g., node ‘h’, ‘i’, and ‘j’ are siblings)
5. Subtree Part of the tree (e.g., Graph with node ‘e’, ‘h’, ‘i’, and ‘j’ are subtree with the root node ‘e’)
6. Leaf: A node without child (e.g., node ‘g’)
7. Internal nodes: A node that has children (e.g., node ‘d’)
8. Degree: Number of a node’s children (e.g., the degree of node ‘b’ is two)
9. Path: Continuous connection of a node (e.g., a-b-e-h is a path from ‘a’ to ‘h’)
10. Level: Number of edges from root to the nodes (e.g., level of node ‘e’ is two)
11. Height/Depth: Max distance from root to a leaf (e.g., the height of tree in Fig. 2.2. is three)

B. Decision Tree

Trees have many applications and usage, such as helping us decipher the simple tasks in our daily lives and even solving complex problems. Trees can be used as basic solutions to many problems with various approach. One of them is a decision tree.

A decision tree used to support decision making by classifying the problems. A decision tree uses the N-ary tree models to represent deceive factors and their consequences. The root is the main problem that we want to solve. Every internal node used as a test case. Every leaf node corresponds to a decision taken as the result for every possible conditions. Every edge corresponds to the possible results of the condition of the node that edge started from. As tree-based algorithm, we will start from the root and follow the compatible branch by value based on the condition then move to the next node. The same thing is applied continuously until we reach the leaf as the final decision.

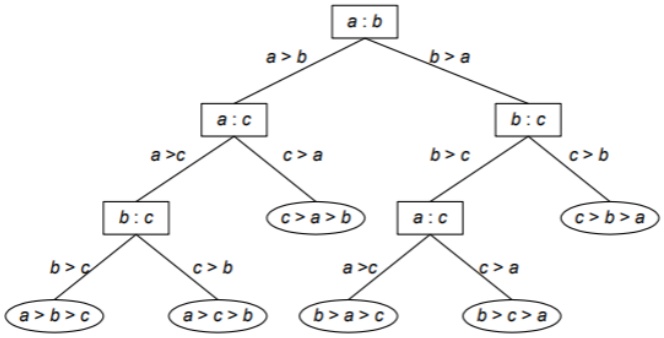


Fig. 2.3. Example of a decision tree to sort 3 elements (https://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2020-2021/Pohon-2020-Bag2.pdf, accessed on December 10, 2021)

As a decision support tool, decision trees have some advantages, which are:

1. Decision made can be easily explained. It make it easier for people to conclude based on certain condition and help people to know strategy to reach their goal
2. Can be used to determine the worst, best, and average values for various scenarios
3. Can be implemented to many applications, such as machine learning

Decision trees surely have disadvantage, they are:

1. Sensitive in the changing of data. In machine learning, it must learn fast to adjust the tree structure caused by changing of data.
2. More complex calculations are required in a huge program and may need to be combined with other algorithms.

C. Decision Tree Algorithm in Machine Learning

Decision tree algorithms as a major part of machine learning support applications in many fields. It can be used to statistically compare data. It can be used in text classification or extraction [2]. Decision trees decide to split a node into two or more sub-nodes use multiple algorithms. The homogeneity of resultant sub-nodes will increase along the creation of sub-nodes. First, the decision tree splits the nodes on all available variables. And then, it will select the most homogeneous sub-nodes after splitting. The type of target variables affects the algorithm selection. Some of the algorithms used in the Decision Tree are as follows:

1. ID3

   To build decision trees, the ID3 algorithm use a top-down greedy search approach through the space of possible branches with no backtracking. Based on the theory that it formulates, it can identify and

predicts the class of an item.

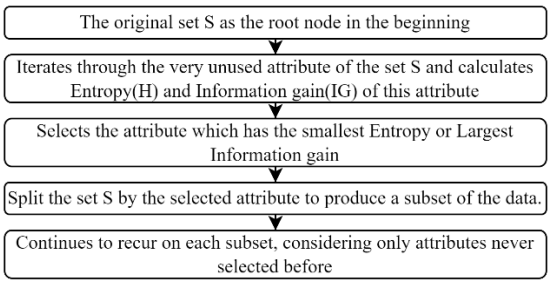


Fig. 2.4. The steps in ID3 algorithm
(Source: Author, modified from literature [3])

2. C4.5
   C4.5 algorithm is an extension to the ID3 algorithm because ID3 have limitation on large values. To overcome this problem, C4.5 algorithm uses a metric called information gain.

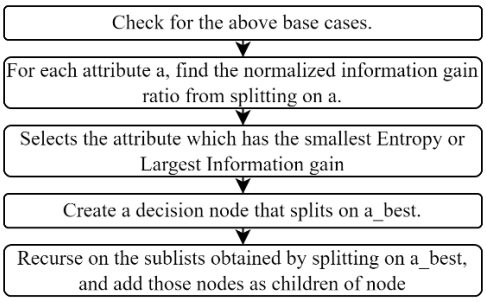


Fig. 2.5. The steps in C4.5 algorithm
(Source: Author, modified from literature [4])

3. CART (Classification and Regression Tree)
   CART explains how an outcome variable's values can be predicted based on other values. A CART output is a decision tree where each fork is a split in a predictor variable and each end node contains a prediction for the outcome variable.

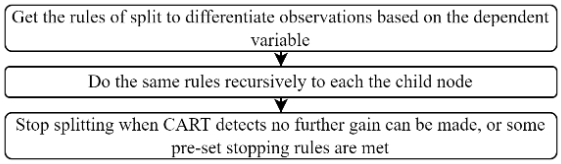


Fig. 2.6. The steps in CART algorithm
(Source: Author, modified from literature [5])

4. CHAID (Chi-Squared Automatic Interaction Detection)
   CHAID builds a predictive model or tree which can determine how variables best merge to explain the outcome in the given dependent variable

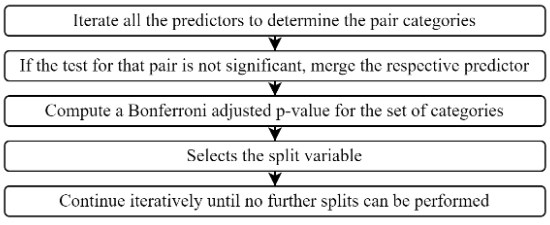


Fig. 2.6. The steps in CHAID algorithm
(Source: Author, modified from literature [5])

III. VIRTUAL ASSISTANT

Intelligent Virtual Assistant (IVA) is an application that utilization information, for example, the user's voice and logical data to give help by noting inquiries in normal dialect, making suggestions, and performing activities [6]. A virtual assistant as special software program works to assist users in daily tasks. At first, virtual assistant can do simple things like answer questions about weather, world time, traffic information, and more. Along with time, they developed into much more complex applications than that. Nowadays, virtual assistants can do many things like calculations, do conversations like human, answer complex questions, perform actions on your smartphone such as call your colleague, and much more. Virtual assistants can recognize their owner's voice and will active when you call them. For example, in Google Assistant you can call it out by saying "Hey Google" or "OK Google" and it will not take long to respond.

Virtual assistants as a bridge of the digital and human world use advanced Artificial Intelligence (AI), Robotic Process Automation (RPA), Natural Language Processing (NLP), and Machine Learning (ML) to extract information and complex data and then process them accordingly to make a decision as instructed. Just like human assistants, virtual assistants should be able to adapt to different circumstances and generate specific patterns to solve problems which need a strong knowledge base. The knowledge base allows virtual assistants to learn from unique and various circumstances. Assistants need training to learn. With a strong knowledge base, a virtual assistant can learn to process the data it receives so that the assistant's knowledge increases and makes it smarter to do more complex things.

Before the virtual assistant can start thinking about appropriate responses to a user's query, it must first determine what the query actually is and understand it. That is where Natural Language Processing (NLP) comes in. Since NLP itself is based on a knowledge base, virtual assistants can read and analyze queries better if there is a more robust knowledge base to work with. NLP comes even after the assistant has fully analyzed the query and obtained an effective response to it. The use of NLP involves NLU and NLG. Natural Language Understanding (NLU) is the process of converting text into structured data that a machine understands. Natural Language Generation (NLG) is the process of transforming structured data into natural language text. Sometimes, detecting the patterns in natural language can be chaotic caused by some challenges such as synonyms/homonyms/slang, missing punctuation, misspelling, or abbreviations. After all, the assistant must determine a suitable way to respond to the user's commands, either by conversation or by an action performed on the smartphone/other connected device.

IV. IMPLEMENTATION AND ANALYSIS

A. Workflow of Virtual Assistants

Virtual assistants are complex programs which require a combination of many algorithms, not just a decision tree. First, virtual assistant needs to receive command from user. In general, there are two forms for inputting commands, either through speech or typed text. Then, the received input will be processed by the machine. The machine will access the database

while sorting the input into structured data so that the words are classified into several parts. This structured data is very important for the machine to decipher what it means. To figure out the user intent, virtual assistant use Pattern Matching Algorithm or NLU.

Pattern Matching Algorithm deployed matching pattern to generate proper response to users' questions, which depends on different types of matching such as simple statements or the meaning of enquiries. Pattern matching is often referred to as "brute force" as the developer of the system needs to describe every pattern and its response [7]. It will look up in a pre-defined dictionary for the corresponding keywords.

The main purpose of NLU is to support human-computer interaction. It is the comprehension of natural human language that allows computers to understand commands without using the formalized syntax of computer languages. NLU algorithms reduce human speech into a structured ontology. So that the data models will consist of semantics and pragmatics definitions. The fundamental concepts of NLU are about intent and entity recognition. Intent recognition identifies the user's sentiment in input text and determine their objective so that it will establishes the meaning of the text. Entity recognition focuses on identifying the entities in a message, then extracting the most important information. There are two types of entities. First is named entities that used to group words into categories such as people, companies, and locations. Second is numeric entities to recognize numbers, currencies, and percentages.

Once the machine understands what the user intents, it has to respond back with natural human language using Neuro Linguistic Programming technique called Natural Language Generation (NLG). Virtual assistant can respond with voice and text, even with an action. Virtual assistant needs to be trained with machine learning so that NLG will generate more natural output like human language.

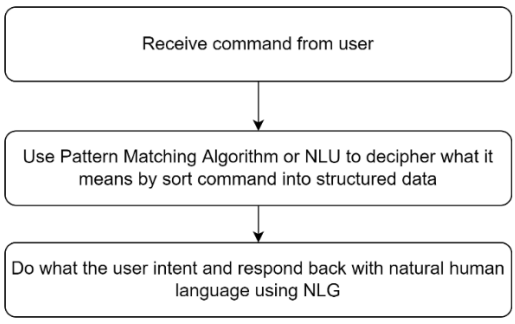


Fig. 4.1 General workflow of virtual assistants
(Source: Author)

B. Speech-to-Text

As previously mentioned, virtual assistant can receive text or speech. If the user commands the virtual assistant by speech, then it must convert user's voice into text first or usually known as Speech Recognition. Speech Recognition making the computer understand what the user speaks. In general program, a computer such that it can hear us and respond us. By 'understand' what the user mean it would convert our voice into suitable format such as Text. Thus, speech recognition is also called as Speech to Text conversion process.

Then, how the implementation of decision tree through this process in the virtual assistant specifically? First, the machine must ensure that the received sound is clear enough so that it would not be misinterpreted (it can distinguish the user's voice and background noise that may be distracting). Machine will turn the sound into wave pattern and then parsing the keywords. It does not need to check whether it is a complete sentence or not. Parsing sound use the technology of Recurrent Neural Networks (RNN). Now audio which is given at input is easy to process, it will be feed into a deep neural network. After feeding small audio chunks of around 20ms to our network it will figure out letter which matches the spoken sound [8].

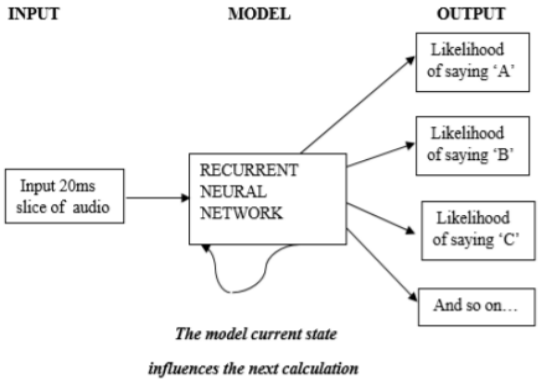


Fig. 4.2. Speech recognition models

It is very possible if there is an error in parsing words, especially on the words that sound similar. To provide optimal results, the machine must be synchronized to the database of languages the user may use. For example, if the user set their virtual assistant with English, then it should prioritize English in deciding the digited text. Figure 4.3 below displays a common decision tree on some virtual assistants in speech-to-text.

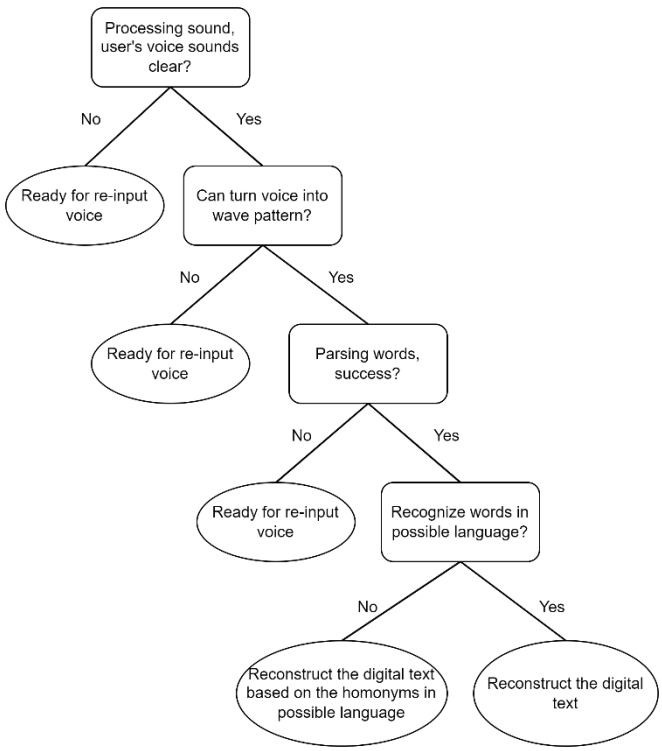


Fig. 4.3. Decision Tree of Speech-to-Text in Virtual Assistant
(Source: Author)

C. Sample Decision Tree of Virtual Assistant: Google Assistant

There are many virtual assistants developed by various companies to achieve different goals. This time, Author will discuss one of the most famous ones, Google Assistant from Google. As we know, Google is a huge company that develops tons of IT products that can work with each other, no exception to Google Assistant. The use of Google assistant is not only limited to its application but can access other applications on the device used by user (e.g., smartphone). Furthermore, Google Assistant also affect other devices that are connected. For example, user can turn on the lamp or maybe the speaker if it has been set up.

The first thing Google Assistant must consider after knowing user's intent is to explore the database in making decision to give appropriate results. In general, there are two sources of google assistant database, they are:

1. Local database, stored on the device and in user's account. Google Assistant gives users the opportunity to choose privacy settings for users by providing easy-to-use built-in controls. Google Assistant can use your data to create more personalized and useful experiences. User can add personal data to google assistant at any time, and this is part of training our assistant to be smarter and know us better so that it will improve in making decision.
2. Global database from the google server. Once again, because Google is a huge company, its applications are integrated with each other, especially Google Search which stores a lot of information in it. This allows Google Assistant to extract various information for consideration in making decisions.

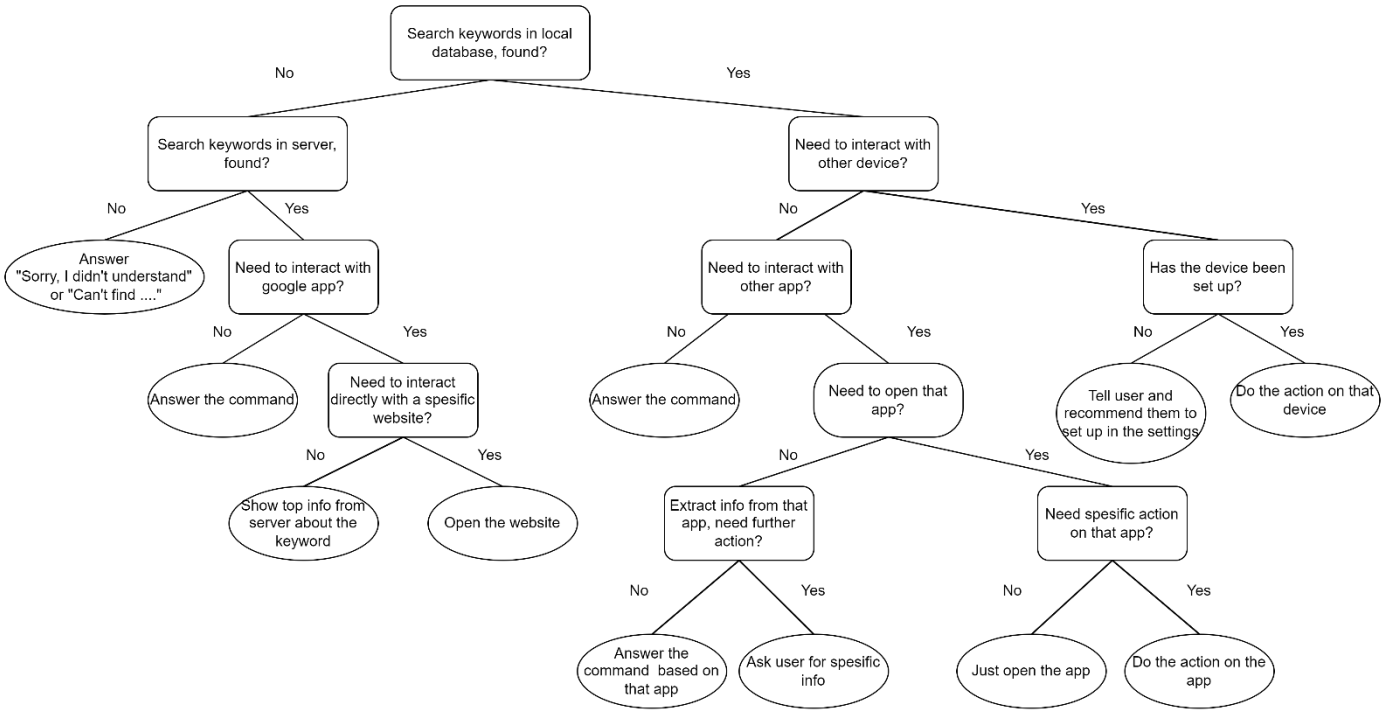


Fig. 4.4. Decision tree of Google Assistant
(Source: Author)

Fig. 4.4. above represents several possible forms of decisions that the virtual assistant makes as response back to our commands. it will depend on the keywords and database sources it uses. It can respond within the google assistant app itself, through other apps, even with actions on other connected devices. Google Assistant also complement its response by speaking with natural human language to increase user experiences. If Google Assistant does not find the data for the keyword, or what the user says/types is not clear, it will say "Sorry, I don't understand" or something like that. This means that the user should try using another keyword.

When google assistant uses a local database, it needs to consider several things regarding the access it needs, they are:

1. If it needs to interact with other devices

   Other devices what is meant here is electronic devices that supported by IoT such as smart lamp, TV, coffee maker, and so on that integrated with Google Assistant. If Google Assistant need to interact with them, then it must check whether that device has been set up or not. If it has been set up, Google Assistant will do the action relate to user's intent. But if it has not been set up yet, or the commands does not match the device, it will tell user and recommend user to set up in the settings. (Examples: "Turn on the lamp", "Turn off the TV")
2. If it just needs to interact within the used device

   The information that Google Assistant can extract comes from related account and applications in the used device. If it has enough information and capability to answer the command, then it can

answer right away. For example, when user asks Google Assistant, “what is my name?”, it will answer according to the provided data by the user. If Google Assistant needs to interact with other app, in this case, Google Assistant’s decision can be grouped into 4 actions based on these conditions:

a. If Google Assistant does not need to open that app and information from user is enough, then it will immediately do the user's command. (Example: ‘Show my photos”)
b. If Google Assistant does not need to open that app but need further information, then it will asks user again until it have enough information to do the task properly. For example, if the user asks, “Send message”, it needs to collecting data about who the user want to message and what the user want to talk about.
c. If Google Assistant just need to open the app, then it will open that app right away. (Example: “Open Calculator”)
d. If Google Assistant need to open the app to do some specific action, then it will immediately open the app and do the action if it has access. (Example: “Play butter by BTS on youtube”)

When Google Assistant does not have enough information from local database, it will consider making decision based on condition in global database. The main source used is the google server. But does not mean that it must interact directly with the Google App. For example, if the user asks, “How is the weather?”, it can explain the weather by extracting information from the server. Google App is the best solution if Google Assistant needs to deal with a lot of information. In this case, there are two conditions, they are:

1. If Google Assistant understands that the user’s intent is only general information

   Google Assistant can access Google search within the app. Then, it can answer right away. For example, if the user asks, “Do you know NLP”, then it will show top info about NLP.
2. If Google Assistant understands that the user’s intent is in a specific website

   Mostly, this case occurs when user literally want to open specific website. However, if user have the corresponding app in the used device, Google Assistant prioritize to open the app instead of the webpage. For example, if the user asks, “Open Netflix” but our device does not have Netflix App, Google Assistant will decide to open Netflix web page.

V. Conclusion

In conclusion, the concept of decision tree is useful in many fields, one of which is in Artificial Intelligence and Machine Learning. One of the utilizations is in decision making by Intelligence Virtual Assistant. A Virtual assistant have many benefits to help humans. The decision tree's structure implemented to classify the conditions to help virtual assistant determine the best to respond back the user. For best performance, virtual assistant also needs robust knowledge base as references. Various algorithms are also combined to support decision making, such as Natural Language Processing which help virtual assistant to understand human language.

VI. Acknowledgment

First and foremost, the Author would like to praise and thank God, the Almighty, who has granted countless blessing, so that that the author has been finally able to accomplish this paper. The author would like to thank all those who have supported the author in the process of writing this paper. Especially, all lecturers of IF2120 Discrete Mathematics Class that has taught and shared their knowledges which are required in writing this paper as final task of this course. The author also would like to thank all the AI and ML engineers who have developed this amazing innovation.

References

[1] Munir, Rinaldi. informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2020-2021/Pohon-2020-Bag1.pdf, accessed on December 10, 2021
[2] Navada, A., Ansari, A. N., & Patil, S. S. *Overview of Use of Decision Tree Algortihms in Machine Learning*. IEEE Control and System Graduate Research Colloqium, 2011, pp. 37-42.
[3] Chauhan, Nagesh. S. www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html, accessed on December 10, 2021
[4] Kotsiantis, S. B. *Supervised Machine Learning: A Review of Classification Techniques*, 2007, pp. 249-268
[5] Ramzai, Juhi. *Simple Guide for Top 2 Types of Decision Trees: CHAID & CART*. towardsdatascience.com, accessed on December 10, 2021
[6] Tulsham, A.S. & Dhage, S. N. *Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa*. Springer Nature Singapore Pte Ltd, 2019, pp. 190-201
[7] Devakunchari, R., Agarwal R., Agarwal, E. *A Survey of Chatbot Design Techniques*. International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-2S2, January 2019
[8] Amberkar, A., Awarsamol, P., Desmukh, G., Dave, P. *Speech Recognition using Recurrent Neural Networks*. IEEE International Conference on Current Trends toward Converging Technologies, Coimbatore, India, 2018

Pernyataan

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 14 Desember 2021

Putri Nurhaliza
13520066