Understanding Human Learning and Decision Making Process based on Jungian Analytical Psychology by incorporating Python and Data Science

By

Nusrat Sumaiya (160021076)

Eshrar Promitee (160021100)

Moumita Khan (160021137)

A Thesis Submitted to the Academic Faculty in Partial Fulfillment of the Requirements for the Degree of

BACHELOR OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING

Department of Electrical and Electronic Engineering

Islamic University of Technology (IUT)
Gazipur, Bangladesh

February 2021

Declaration of Authorship

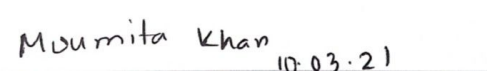
We, Nusrat Sumaiya (160021076), Eshrar Promitee (160021100), and Moumita Khan (160021137) declare that this thesis is titled, 'Understanding Human Learning and Decision-Making Process based on Jungian Analytical Psychology by incorporating Python and Data Science' and the works presented here contains our original work. We confirm that:

- This thesis was completed in partial fulfillment of this university's Bachelor of Science in Electrical and Computer Engineering degree.
- This work has not been applied to any other university for the purpose of receiving a degree.
- We have always simply attributed the sources when we have consulted other people's written work.

Submitted by:

Nusrat Sumaiya (160021076)

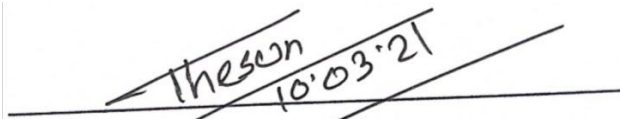
Eshrar Promitee (160021100)



Moumita Khan (160021137)

Understanding Human Learning and Decision Making Process based on Jungian Analytical Psychology by incorporating Python and Data Science

Approved by:



Md. Thesun Al-Amin

Supervisor and Assistant Professor
Department of Electrical and Electronic Engineering
Islamic University of Technology (IUT)
Board Bazar, Gazipur-1704

Date__ 10·03·2021

Abstract

An updated version of MBTI is the Jungian Analytical Archetypes where the same 16 types of personality exist but the typecasting is done not by the letters but the cognitive functions and their positions in respective slots. It is an analytical psychology psychotherapeutic technique in which the analyst and patient Collaborate to put unconscious psyche components into a more balanced relationship with conscious knowledge and experience. In our study, a Jungian analysis-based customized questionnaire was made and asked to people. Both supervised and unsupervised approaches to machine learning were used in our study. The Random Forest Classifier and k-means clustering algorithms were used to generate predictions of personalities with a supervised learning module. The results of this study's methodology were compared to those of other existing methods, and the results showed that this methodology performed better in terms of accuracy and reliability. Our study also has a great potential of helping people in various ways.

Acknowledgments

Firstly we would like to offer our gratitude to Almighty Allah (SWT) that we could complete our work without any mishaps and with good health.

We are grateful to our thesis supervisor, Md Thesun Al-Amin for the guidance and the support he provided during our tough times. Without his support, this project could never be completed smoothly and on time. He also helped us to attain the knowledge that was necessary to finish the work we started. We are thankful to him.

We are also very thankful to the teachers of the department of EEE, IUT for their guidance throughout our work.

Finally, we offer our gratitude to our caring and loving families and friends who constantly supported and inspired us throughout the journey.

Contents

List of figures

List of Tables

Abbreviations

Te- Extraverted Thinking

Ti- Introverted Thinking

Fe- Extraverted Feeling

Fi- Introverted Feeling

Ne- Extraverted Intuition

Ni- Introverted Intuition

Se- Extraverted Sensing

Si- Introverted Sensing

I- Introverted

E- Extraverted

N- Intuitive

S -Sensing

T- Thinking

F- Feeling

J- Judging

P- Perceiving

AI- Artificial Intelligence

ML- Machine Learning

NLP- Neuro-Linguistic Programming

MBTI- Myers–Briggs Type Indicator

RP- Random Partition

3D- Three dimensional

WCSS- Within Cluster Sum of Squares

SVC- Support Vector Classifier

k-NN- k-Nearest Neighbours

Chapter 1

Introduction and background

1.1 Overview of Jungian Analysis

The human mind is always a mystery. What motivates them, what they are usually interested in, how they take decisions, what do they prioritize when making decisions, what are their learning methods—specialists are trying to figure these out in-depth. Depth psychology or Jungian Analytical Psychology is an effective method for understanding these. This field of science focuses on behavioral patterns and how it affects the ego and the subconscious and the unconscious and forth to make decisions.

Carl Jung was the first to observe cognitive senses, which later was termed as cognitive functions. [17] He also theorized about the top two cognitive function slots which are the primary motivation for individuals to navigate life.

Katharine Cook Briggs and Isabel Briggs Myers developed Myers Briggs Type Indicator (MBTI) based on Jung's theory on psychoanalysis. [17] Later Dr. John Beebe and Dr. Linda Berens contributed tremendously to the improvement of these theories. [1] [2] [7] [8] [11] In total there are 16 Archetypes among human beings. They are also known as Personality Types.

This field of science is very metaphysical in nature. In other words, it is not a tangible output or concept to grasp. Since it deals with the four sides of the mind, it has a lot to do with abstraction.

Nowadays, MBTI has become a topic of discussion in popular culture. It is a type of self-report questionnaire which reveals how an individual perceives information and makes decisions based on them. It focuses on typing as I (Introversion) vs E (Extraversion), N (Intuition) vs S (Sensing), T (Thinking) vs F (Feeling), and J (Judging) vs P (Perceiving).

This personality assessment has its flaws and limitations. Because it focuses on typing through letters and giving percentages on those said letters. A person can be both Feeling Type and Thinking Type. The J vs P statement is also flawed because it is not about order and progress literally but has a lot to do with cognitive functions. So, the most accurate method is to typecast people based on their cognitive functions or cognitive senses as Carl Jung implied.

This field of study has so many possibilities and room for improvement and application. Having a proper understanding of cognitive functions will open doors for so many opportunities. Children of the lesser-known archetypes will not be understood and mistreated. Because certain types have certain methods of learning and understanding concepts, there can be 16 different curricula to benefit future generations. Besides that how human nature works and what is the source of different cognitive-behavioral patterns, whether it is genetic or not, that can be figured out with the help of Brain-Computer interface data.

1.2 Literature review

It's been observed that the archetype of an individual can be deduced better from the cognitive functions they use rather than the letters mentioned in the type indicator. That's the basis of Jungian analysis. Carl Jung first observed a pattern within human behavior that got more accurate and updated throughout the years.

In Jungian Analysis, there are the same 16 types but the typecasting is done based on the cognitive functions and in which slot they are placed.

The 16 types are expressed with four letters. Introvert or Extrovert (I vs E), Intuitive or Sensing (N vs S), Thinking or Feeling (T vs F), Judging vs Perceiving (J vs P).

The 16 Personality Types are—ESTJ (Overseer), ESTP (Persuader), ENTJ (Chief), ENFJ (Mentor), ESFJ (Supporter), ESFP (Entertainer), ENTP (Visionary), ENFP (Advocate), ISTJ (Examiner), ISTP (Craftsman), INTJ (Strategist), INFJ (Sage), ISFJ (Defender), ISFP (Artist), INTP (Engineer), INFP (Dreamer). These types have many different names among different communities. [12]

What is prevalent in the online community is that the tests are based on the 4 letter dichotomy like I vs E, N vs S, etc. Such methods do not produce accurate results.

Our proposal is about conducting the test by asking certain questions and figuring out each participant's types based on the answer as these answers focus more on their cognitive features.

The questions are based on various cognitive functions, social interaction styles, and temperaments. According to the Type Grid [12] [11] [7], it is very easy to figure out the archetypes of each individual. The answers to the questions asked indicate different components of the Type Grid.

This method gives a better understanding of human nature.

The entire study was made to have an overview of data science over psychology. Our study refers to two particular approaches of machine learning which are Supervised and Unsupervised Machine Learning. Random Forest Classifier had been used to train the data set and execute necessary tests to provide the results covering the supervised data scheme. On the other hand, the K-means clustering algorithm was used to implement the unsupervised machine learning scheme over the data set that was created for the study.

1.3 Thesis objective

The main objective of our thesis is Understanding Human Learning and Decision Making Process based on Jungian Analytical Psychology by incorporating Python and Data Science. Specific objectives of our thesis are as follows:

- To extract information from our collected data and determine personality types from that. This incorporates writing surveys, collection of data, development of optimization calculations, etc.
- To have defined numbers of clusters for respective personality types and their traits using k-means clustering algorithm.
- To obtain the relations between the mutually exclusive parameters using Pearson correlation mapping.
- To summarize future possibilities and conclusions from obtained results and discuss the areas of potential applications.

Chapter 2

The 8 Function Model

2.1 Cognitive Functions, John Beebe's 8 Function Model

Cognitive functions are prevalent in every aspect of our lives. Starting from our behavior and reactions to our surroundings, to what we prioritize, how we affect one another as individuals and collectively, our interests and what drives us, what annoys us, and also how we interact with one another. These functions are a part of human nature, not nurture. [1] [2] [8] [12] First, Carl Jung observed cognitive patterns in human behavior. [17] Later, Dr. John Beebe updated his work by proposing an 8 Function Model. [1] [2] [8]

Extraverted functions:
Extraverted functions are outward-focused. Those who are primary users of these functions, in other words, Extraverted functions hero users (Exxx types), tend to focus on their surroundings to make decisions. When a type is called a user of any function, this means their ego side uses this particular function. [1] [2] [8] [12]

Introverted functions:
Introverted functions are source functions. They are focused inwardly. The Introverted functions users, who have Introverted functions in the hero slot (Ixxx types) tend to focus on internal processing. [1] [2] [8] [12]

Perceiving functions:
Perceiving functions are used to gather data. It can be present-focused, past-focused, or future-focused. That depends on the different users of these different Perceiving functions. Perceiving function hero users are ExxP and IxxJ types. [1] [2] [8] [12]

Judging functions:
Judging functions, as the name suggests are used for making decisions. It could be focusing on collective or inward values or beliefs or logic of an individual. Judging function hero users are ExxJ and IxxP types. [1] [2] [8] [12]

The Judging functions are–

- ❑ *Extraverted Thinking, Te*

- *Introverted Thinking, Ti*
- *Extraverted Feeling, Fe*
- *Introverted Feeling, Fi*

Extraverted Thinking, Te:

Like any other Extraverted function, Te emphasizes collective data. A Te user can easily tell what others are thinking. This may not necessarily be true literally, more like it depends on certain situations and how efficient the user of the certain personality type is. This is the "thinking" of a crowd, the rationale. In other words, Te is the empirical evidence and statistical probability of something being true. Te users tend to have reference points for their thinking and understanding. If Ti is the logic (inward thinking) of an individual, Te on the other hand is the beliefs of a society. Te users value perception and have the ability to manipulate perceptions of others. They also are good at organizing more than others. One thing to be aware of is that these functions come in axes. This means the axis works like Ti-Fe, Te-Fi, etc. Te users have tendencies to prioritize credentials and credibility more than Ti users. If a user has immature Te, they can be narcissistic. They are also prone to have confirmation bias and superiority complex. The Te user types are xxTJ, xxFP (Delta and Gamma Quadra). [1] [2] [8] [12]

Introverted Thinking, Ti:

Ti is in simple terms, the logic of an individual. Ti users prioritize one's thinking rather than what the beliefs are. Where Te users focus on statistical probability, Ti users verify it as right or wrong with their understanding. Kind of like true-false statement verification. They value facts and truth rather than perceptions. If Te is inductive reasoning, Ti is deductive reasoning. Since this is an inward function, Ti users exhibit traits like thoughtfulness and self-assuredness. At the same time, they suffer from self-doubt and self-deprecation, and arrogance (for the case of immature Ti users). They also tend to be blunt and literal at times (IxFJ). The Ti users are xxTP, xxFJ (Alpha and Beta Quadra). [1] [2] [8] [12]

Extraverted Feeling, Fe:

Fe is the collective form of feelings and values, in other words, ethics. An Fe user can feel what others can feel in real-time. This can be compared to empathy as well as reflecting on other's emotions. A scenario can be illustrated if there is a room full of Fe users, if a sad person enters the said room, then all the users have a greater chance to become sad as well. Fe users prefer to maintain harmony. Since functions come as axes and pairs, Fe is paired with Ti. This means Fe users are also Ti users. In the cases of Fe and Ti users, they focus on inward thinking and outward feeling. Fe users are tuned with the emotions of others, but they have a hard time figuring out their own emotions. Some of the types are also unaware of their feelings. The Fe users are xxTP, xxFJ (Alpha and Beta Quadra). [1] [2] [8] [12]

Introverted Feeling, Fi:

Fi means the morals of an individual. Fi users are tuned with their own emotions but have a hard time understanding other's emotions. They are more of a sympath as they relate to

other's emotions. They tend to look out for themselves and their self-interests. It is not like they are completely unaware of what others feel, they can tell if it matches with their own experiences (past knowledge). Fi users decide actions as good or bad, where Fe users (due to their Ti) judge as right or wrong. Mature Fi users usually have a high moral compass. There's a misconception of calling Fi users selfish but that is not true. Selfishness has very little to do with cognitive functions. The Fi users are xxTJ, xxFP (Delta and Gamma Quadra). [1] [2] [8] [12]

The Perceiving functions are–

- ❑ *Extraverted Intuition, Ne*
- ❑ *Introverted Intuition, Ni*
- ❑ *Extraverted Sensing, Se*
- ❑ *Introverted Sensing, Si*

Extraverted Intuition, Ne:

Ne users are aware of multiple outcomes. They are oriented with metaphysics, the what-ifs for all future. They can be called visionaries. These types are aware of what others do and the intentions of others. Also, since they can see multiple outcomes, they're able to make plans A to Z for anything. Ne users are sometimes chaotic (especially xNxP) and they create chaos to make progress in work. Ne users can pick multiple paths, see patterns very easily but they have a hard time choosing the best possible path. This is why it is common for Ne users to have troubles with what they want in life. The Ne function is paired with Si, so the axis is Ne-Si or Si-Ne. Ne users tend to make decisions through prescience. The Ne users are xNxP, xSxJ (Alpha and Delta Quadra). [1] [2] [8] [12]

Introverted Intuition, Ni:

Ni is more of a self-oriented function. Ni users are aware of the what-ifs as well but for their future. It's sort of a tunnel vision type of function, where the users can choose the best possible path to take decisions. Ni users value personal freedom over everything. These types of people move forward through sheer willpower. They usually don't have much discipline since they're mostly about wanting rather than responsibility. Sometimes people confuse the two terms, willpower and discipline. These are two different things. Ni is paired with Se in the Ni-Se and Se-Ni. The Se users are xNxJ, xSxP (Beta and Gamma Quadra). [1] [2] [8] [12]

Extraverted Sensing, Se:

Se can be associated with physics and mechanics. Se users tend to focus on the moment, the present. They have the mechanical aptitude more than any other users as manipulating the physical realm comes naturally and easily to them. They are forgetful kinds of people. Se means short-term memory. They exhibit impulsiveness, reactiveness, and spontaneity. Se users love to give others a good experience. Immature Se users can be domineering, overbearing, and overwhelming. Se users have a habit of storing memories in objects (somewhat of totems). As paranormal as it sounds, it is close to the quality of psychometry. It

can also be called volatile memory, in that case, Se is similar to a RAM where new information enters and pushes the older information away. Se is paired with Ni in the axis of Se-Ni or Ni-Se. The Se users are xNxJ, xSxP (Beta and Gamma Quadra). [1] [2] [8] [12]

Introverted Sensing, Si:

Si users are past-focused. They are very dutiful, disciplined, loyal. They value and take pride in their beliefs, traditions, and conviction. They are very safety-oriented and prefer comfort. Si users love to receive a good experience. Si users have long-term memory. In simpler terms, they are like the library of Alexandria, an archive where they store past data. If Se is a RAM, then Si is synonymous with a hard drive disk. The Si augmented with Ne in the Ne-Si or Si-Ne axis helps these types to predict. That is how they learn through prescience. Since Si users do not have Se in their ego, they take longer to react to something. Unless they have already experienced something similar, where they use their foreknowledge to adapt to those situations. The higher Si users tend to protect things which is why they are called the Guardian temperament. The Si users are xNxP, xSxJ (Alpha and Delta Quadra). [1] [2] [8] [12]

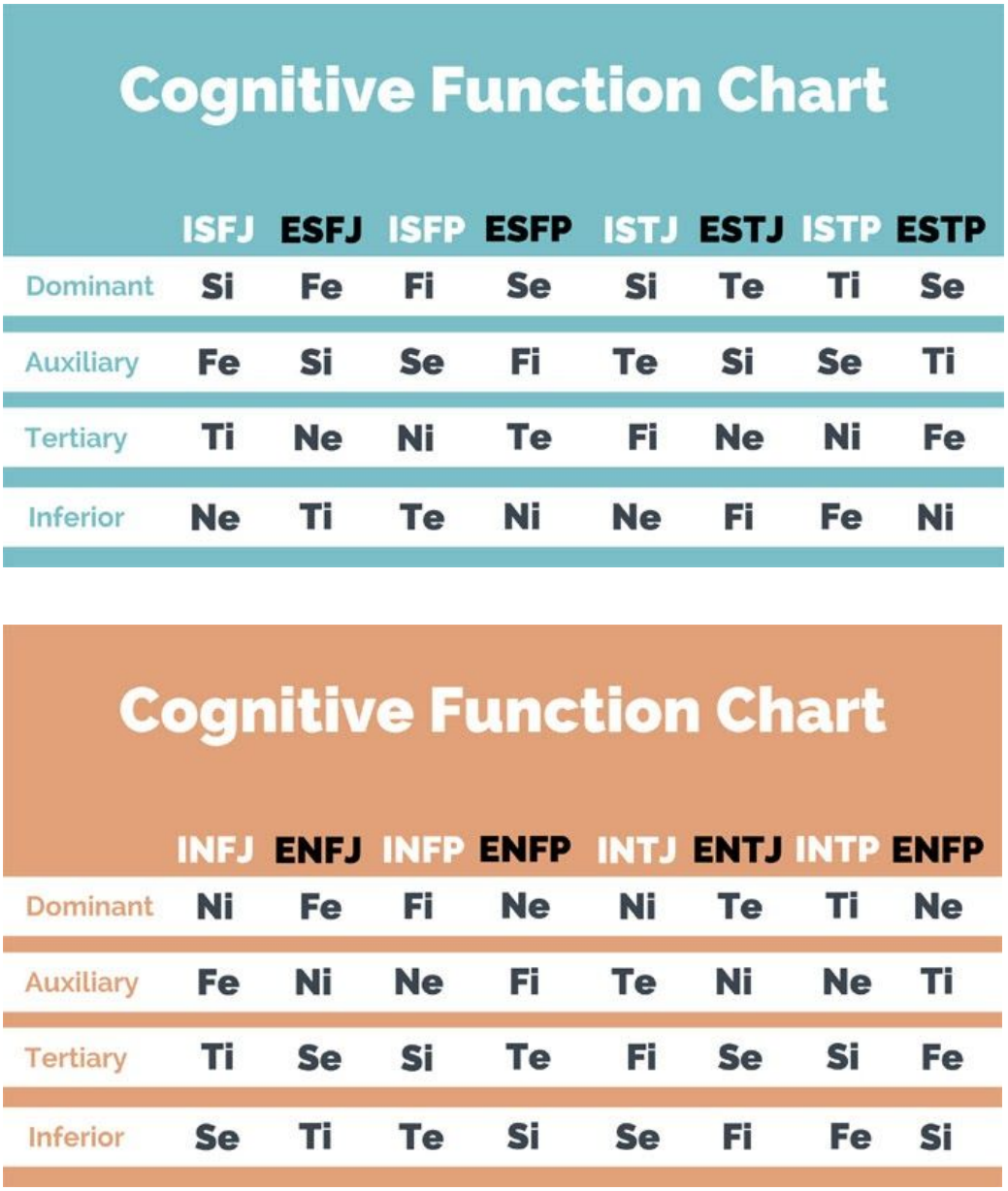


Fig. 2.1 The Cognitive Function Charts. [13] [14]

2.2 Quadra

Through cognitive axis analysis vectors are originated and four types of quadras start functioning. Quadra is like an output vector of the cognitive functions axes.

The four Quadras are—

- ❑ *Alpha Quadra*
- ❑ *Beta Quadra*
- ❑ *Gamma Quadra*
- ❑ *Delta Quadra*

Alpha Quadra:

They are called crusaders. Crusaders have the axis Fe-Ti/Ti-Fe and Si-Ne/Ne-Si. The types in this Quadra are ESFJ, ENTP, ISFJ, INTP. Since these types have the same cognitive functions, they have a similar pattern of thinking and gathering information. But since each of these types have a different order of function stacking, they prioritize different things. Crusaders are all about justice and fairness and honor. They prefer to stay in their comfort zone. That is because they are systematic and also they all have Si. Systematic also means that these types prefer doing a task in the best possible way which will bring the best possible outcome. The Alpha quadra types are all informative types. They tend to be ambiguous and vague with their words because they prefer to keep the options open for others to interpret. These types are very dedicated to making others happy but have a hard time accepting their happiness. If the justice of these types is not met, their sense of justice gets corrupted with selfishness and it turns to vengeance and vindictiveness. Alpha quadra has the highest social compatibility with the gamma quadra and then the beta quadra. [9]

A flowchart was generated by Code2flows to visualize the logical stream of Quadra and Archetype origination.

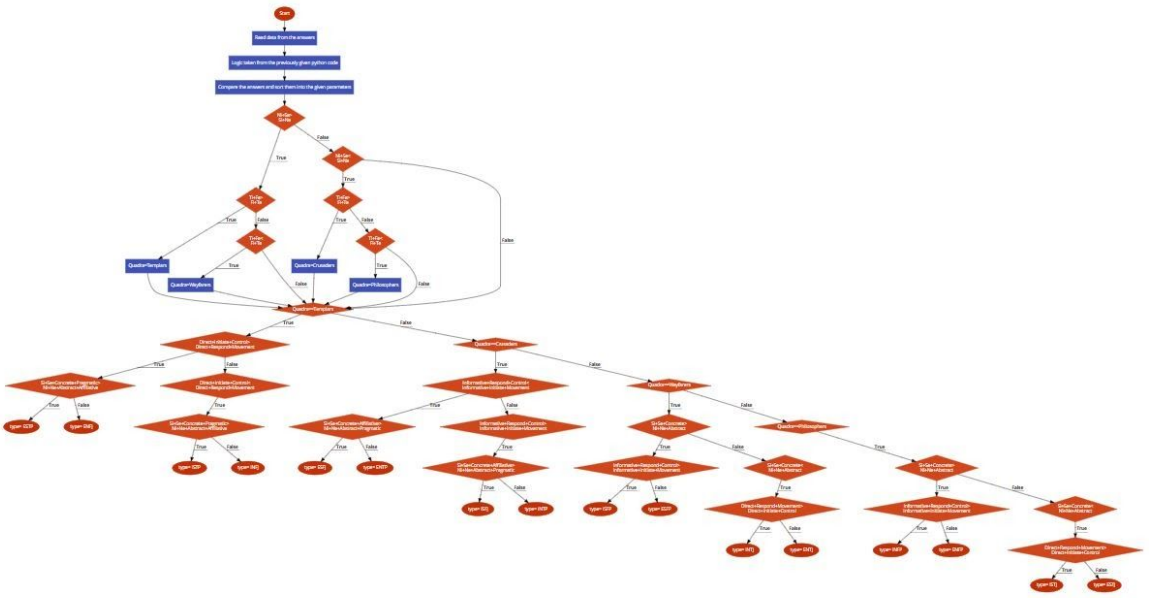


Fig. 2.2 Flowchart to determine the Quadras.

Beta Quadra:

They are called Templars. Templars have the axis Fe-Ti/Ti-Fe and Se-Ni/Ni-Se. The types in this Quadra are ESTP, ENFJ, ISTP, INFJ. Templars seek to make people better human beings. They love to mentor and counsel and heal. They value personal freedom since they all

are Ni users. All of the Templars are interest-based so they are likely to choose happiness over comfort when making decisions. They are directive kind. So they mean what they say and they say what they mean. Templars are more likely to be a leader than a follower. These types love to give others a good experience. These types have the highest capacity of forgiving others but at the same time, they will ghost someone if they are not being responsible. Beta quadra has the highest social compatibility with the delta quadra. [9]

Delta Quadra:

These types are called Philosophers. They have the axis Fi-Te/Te-Fi and Si-Ne/Ne-Si. The types in this Quadra are ESTJ, ENFP, ISTJ, INFP. The Delta Quadra usually chooses fields related to academics and politics. They are also seen performing civic duties. Philosophers are great archivists with their Si and Te. Since they have Fi and Si, they choose to look out for themselves before others. Self-comfort, reputation, and faith, beliefs, prestige are very important to these types. They all are affiliative and so they prefer to do the right thing. This quality also gives them tremendous power to change the beliefs and perceptions of a society or a race. The historians are mainly of the Philosopher Quadra. They seek and explore the truth that serves their Si and Te. Corrupted beliefs can lead to these types becoming selfish and aristocratic. [9]

Gamma Quadra:

These types are called Wayfarers. They have the axis Fi-Te/Te-Fi and Se-Ni/Ni-Se. The types in this Quadra are ENTJ, ESFP, INTJ, ISFP. Like Templars, these types also value personal freedom a lot. But they use freedom mainly for personal achievements. They are competitive, passionate, driven, independent, and realistic. Since all the types are pragmatic, they prefer to be individualistic. They will do what works, what is more practical, and what will bring an effective result in the fastest way. Like every other quadra, they also have the potential to be corrupted. The more corrupted they are, the more shallow they become where they are prone to snatching away what is others'. Wayfarers love to travel more than any other type of quadra. They excel in art and engineering, business, and leadership. [9]

2.3 Function slots and their cognitive attributes

Caul Jung called it cognitive senses. Cognitive functions behave differently when they are placed in different stacks or orders. There are 8 slots for cognitive attitudes. They are—

- ❑ *Hero function*
- ❑ *Parent function*
- ❑ *Child function*
- ❑ *Inferior function*
- ❑ *Nemesis function*
- ❑ *Critic function*
- ❑ *Trickster function*
- ❑ *Demon function*

Hero function:

The hero function is the gateway to the ego side of the mind. With the functions placed in this slot, the user has no trouble using them. In other words, the user can easily use the hero function where no mental energy will be used. If a type is Ni hero, that means they will have no trouble thinking about what they want. A mature Ni hero can easily pick the best possible outcome for their situations. A mature Si user has no trouble being responsible. It is an optimistic function. The odd-numbered slots are optimistic whereas the even-numbered slots are pessimistic. Hero slot is also known as a dominant or primary slot. [1] [2] [8] [12]

Parent function:

Any of the cognitive functions in the parent function slot shows that the user is very responsible with that function. If a type is Fe parent, that means they are very ethical and they are responsible about it. If a type is Ni parent, then that means they are responsible about their future visions and diligently work to achieve them. Even though it is a pessimistic slot, the 16 types do not have to spend that much mental energy for using that function because the responsible nature is inherent. The parent slot is also known as the auxiliary or the secondary slot. [1] [2] [8] [12]

Child function:

It is also called the tertiary function. Since it's called child, any type has trouble using the functions in this slot because it's not that stable. Mastering their child function, each type needs to practice a lot. One of the positive sides of the child slot is it's very unadulterated and pure. If a user is a Ti child, that means making decisions based on logic will not come that easily but their logic is one of the best kinds as it's unadulterated. If a type is Ni child, that means they have a sort of clear idea of the what-ifs and what they want but since their parent is Se, so they have sometimes trouble to work according to what they want or figuring out what they want. This one is an optimistic function. [1] [2] [8] [12]

Inferior function/ Aspirational:

This slot is where all the types' fears lie. At the same time, they are aspirational about it in their subconscious. Every type has certain anxiety and the inferior function reflects their anxiety. They are pretty efficient using that function but the highest mental energy is spent when the users use their inferior function. A proper amount of practice fixes this problem and makes the function aspirational instead. This sort of behavior is the healthy gateway to the subconscious side of the mind from the ego side. If a type has Ne inferior, they will always think that other people have bad intentions. It is a pessimistic function. [1] [2] [8] [12]

Nemesis function:

This is the first function in the shadow side of the mind. Contrary to popular beliefs, shadow does not oppose ego. It is more of a complement. Like the opposite of Ni is not Ne, it is Se. So an Ni hero has an Ne nemesis. The hero function can also transform into nemesis and vice versa and that happens due to cognitive transition. But generally speaking, the Nemesis

function is something that makes an archetype worry. If a type is Fe hero, then they are Fi nemesis. An Fe hero is about empathy and harmony but due to their Fi nemesis, they think they are a bad person. Even though it is about the worry of an archetype, this slot is still optimistic. This slot is also the gateway to the shadow side of the mind. [1] [2] [8] [12]

Critic function:

This slot opposes and also complements the parent slot. It can be at times destructive because it is in the shadow side as well as it is a pessimistic slot. The critic function as its name suggests is a critic, most of the time, an unhealthy one. An Fi critic will be more of a harsher and more destructive function than an Fi nemesis. Because it will criticize one's morals. [1] [2] [8] [12]

Trickster function:

The archetypes are unaware of any function in the trickster slot. If an archetype is Te trickster then they are unaware of the beliefs and rationale, they are also unaware of what others think. In the same manner, an Fe trickster is unaware of what others feel. For the inward functions, the introverted functions, if they are in the trickster slot, then in the case of a Ni trickster, the archetype generally will be unaware of what they want. Comparing it with the child slot, if an archetype is Ne child so naturally they will be Ni trickster. Ne child is busy being aware of what everyone else wants, so they do not think much about what they want in general. [1] [2] [8] [12]

Demon function:

The last slot in the shadow side is also the gateway to the superego side of the mind. It is the lowest awareness slot in the human mind. Demon functions are paired with the inferior functions. Vulnerabilities lie in inferior function. So if the inferior function gets affected, this leads to a Faustian bargain. Demon function is the gateway to the superfluous ego of the mind. Demon function may usually lead to the peril of an individual but if treated and nurtured well, it can also lead to the enlightenment of a human being. If a Te user is in the demon slot, it usually results in acts that make sure that no one has a voice or opinion to give. Because (ExFJ) Ti inferior wants to be listened to. If no one listens to these types, Te demon awakens and makes sure nobody gets listened to. Like if their voice does not matter then nobody else's does either. [1] [2] [8] [12]

2.4 How to name the types based on the cognitive function order

There are some rules on how the functions are placed in the function stack. They are—

- ❏ The first and the fourth slots (hero and inferior) and the second and the third slots (parent and child) will be in axis pairs. If the hero is Ni, then the inferior is Se. If the parent is Fe then the child will be Ti. This tells us that hero/ inferior is a perceiving

pair, then the parent/ child is a judging pair and vice versa. This is the case for the ego-conscious side.

- ❏ If the first function is extroverted then the second and the fourth function are naturally introverted and vice versa. This is why after Ni, there will be Te or Fe and the inferior will be Se. Anything other than that is not possible. Similarly, the third child function will be introverted (Ti or Fi) when the hero function is introverted.
- ❏ The subconscious side has the same functions but in the opposite order. Considering the previous example where the stack was Ni-Fe-Ti-Se, the subconscious will be Se-Ti-Fe-Ni.
- ❏ The shadow or the unconscious side has the introverted or the extraverted side of the exact functions given in the same order. For the same previous example, if ego is Ni-Fe-Ti-Se, the shadow is Ne-Fe-Te-Si (nemesis-critic-trickster-demon).
- ❏ For the superfluous ego, the superego, the shadow functions will be in the opposite order. So if ego is Ni-Fe-Ti-Se, superego will be Si-Te-Fi-Ne. It can also be seen as the opposition (as extraversion or introversion) of the function in the subconscious side. [9] [18]

To name a type from its functions or vice versa, there are a few rules to that as well. They are—

- ❏ The previous example will be used again. The first function shows whether the archetype is introvert or extrovert (I vs E). Here the hero function is Ni, so the first letter of the archetype is I. Assuming the type to be Ixxx for now.
- ❏ The Hero function in this stack is Ni. This means the type has an awareness for intuition-related perceiving function over a sensing one. So naturally, the type is INxx instead of ISxx.
- ❏ The second function in this stack is Fe. This means the type has an awareness for a feeling-related judging function over a thinking one. So naturally, the type is INFx instead of INTx.
- ❏ The last letter placement (J or P) follows a rather unorthodox rule. The letter J or P does not tell whether a person is judgmental or perceiving or not (all types can be judgmental or not depending on how they are nurtured). J vs P is deducted based on whether the archetype's top extraverted function is Judging or Perceiving kind. For Ni-Fe-Ti-Se, the first extraverted function Fe is a judging function. So the type is INFJ. So the misconception needs to be corrected, INFJs are more perceiving type (because hero function Ni is a perceiving function) even though they have the letter J in their name.
- ❏ In the same way, it is observed from the functions, that the subconscious of INFJ is ESTP, the shadow is ENFP and the superego is ISTJ. [9] [18]

The functions can be derived in the same manner from the letters. This can be illustrated with another example.

- ❏ For ESTJ, the first function will be extraverted. Since there are letters S and T, then the possible top two functions are Se, Si, Te, Ti. They can be any order.
- ❏ But the J tells that the first extraverted function is a judging function. Among the possible cognitive functions, the only Extraverted judging function is Te. So the Hero is Te. Naturally the inferior is Fi.
- ❏ The parent function will be introverted. Since the second letter is S, so the only possible introverted parent function is Si. So the child function will be Ne.
- ❏ The ego side function stack of ESTJ is Te-Si-Ne-Fi. Subconscious side, Fi-Ne-Si-Te (INFP). Shadow side, Ti-Se-Ni-Fe (ISTP). Superego side, Fe-Ni-Se-Ti (ENFJ). [9] [18]

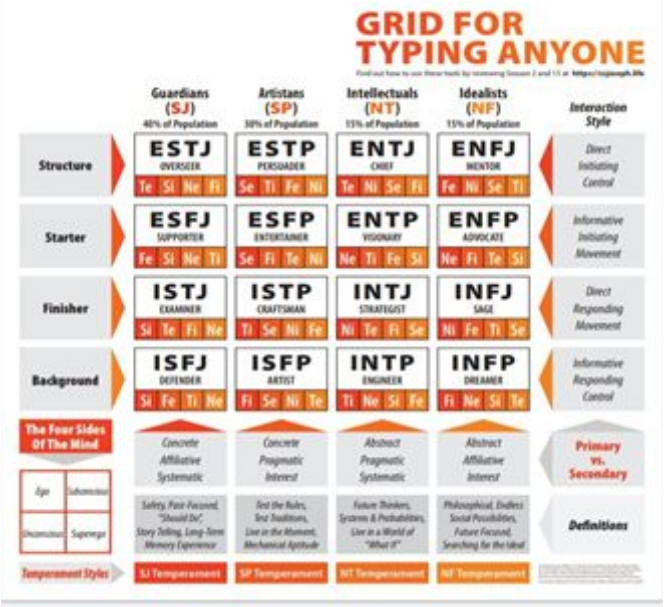


Fig. 2.3 Archetype Typing Grid. [12]

2.5 Four sides of the mind

Human mind works in different ways. Every human has four sides and in different stages of life, those sides come to surface.

The four sides of the mind are—

- ❏ *Ego*
- ❏ *Subconscious*
- ❏ *Unconscious*
- ❏ *Superego*

Ego/ Conscious:

The Ego side is where one's conscience is. Every type is aware of their conscious ego. The Hero function is the gateway of the Ego side. Ego consists of Hero, Parent, Child, Inferior/

Aspirational. It is the primary side of the human mind. That means everyone feels most comfortable using this side of the mind. The functions in the Ego side have the most mental awareness. That is, these particular functions in the four slots take the least mental energy to use them, Hero being at the top. [1] [2] [8] [10]

Subconscious:

The functions in the subconscious side are the flipped stack of the hero side. As explained in the earlier section, if the Ego side is ENTJ (Te-Ni-Se-Fi), then Subconscious will be ISFP (Fi-Se-Ni-Te) where the functions are flipped, and the letters will be opposite as well. The Inferior function will become Aspirational and act as a gateway to the Subconscious side. It could also be called the Secondary side of the human mind. The subconscious side usually is about how a type wants to see themselves. An ENTJ would want to have strong morals and also would love to turn art into a form of business. They always have something to show and their subconscious would help them achieve it. People generally live vicariously through others with their Subconscious side. If a type has a gradual healthy progression to this side, it can be very beneficial in all ways, where the Inferior turns to an Aspirational function. If the progression is unhealthy then it could be bad for oneself and the people surrounding that person. [1] [2] [8] [10]

Shadow/ Unconscious:

The functions in the shadow side are in the order, Nemesis, Critic, Trickster, Demon. For ENTJ (Te-Ni-Se-Fi), the Shadow will be INTP (Ti-Ne-Si-Fe). Here the first and the last letters are different. The gateway of the Shadow side is the Nemesis function. The Shadow is the unconscious side of the mind. It is the side one is unaware of. It could be somewhat of another Secondary or Tertiary side of the mind. If there is a healthy positive gradual shifting of the sides, then the Nemesis function turns into an Ally function. The Critic function changes into something where one's wisdom lies. The Trickster evolves to Master and the Demon becomes an Angelic function. But if the progression is not treated well, the cognitive transition will lead to a kind of Faustian Bargain. [1] [2] [8] [10]

Superego:

The last side is not usually taken lightly. Because it has the potential to lead to one's destruction. On this side, the functions of the Shadow side are flipped. So, the Superego of ENTJ (Te-Ni-Se-Fi) will be ESFJ (Fe-Si-Ne-Ti). The DEmon function is the gateway of this side. In popular culture, the best portrayal of the Superego was done in Christopher Nolan's movie "The Dark Knight" character the Joker. Joker was an ENTP initially, but when he could not navigate his life in his Conscious, Subconscious, or Unconscious state, he entered the Superfluous Ego (ESFP) and became chaotic evil with his Se demon. The only good thing about this side is it eventually leads to enlightenment. How a forest gets destroyed by a storm and later gets reborn, the same can be said about human minds. Of course, it does not always occur. Sometimes one may enter the Superego side in a structured way. That is when the functions become Ally, Wisdom, Master, and Angelic. People become the best version of themselves. [1] [2] [8] [10]

Chapter 3

Interaction Styles, Temperaments, and The 16 Archetypes

3.1 Interaction Styles

There are different communication styles in social situations. Each style has different components. Dr. Linda Berens explained these components and style in her works for years. Later, she created a Type Grid to make accurate typecasting of the 16 personalities. [7] [11] Later, C.S. Joseph renamed the terms for understanding them easily. [9] [12] These components are—Directive, Informative, Control, Movement, Initiate, Respond.

3.1.1 Directive vs Informative

Directive people's speech is concise and leaves no room for further interpretation. They mean what they say and they say what they mean. Informative people, on the contrary, leave statements that are open for different interpretations. An illustrated scenario of the two kinds can be—if a task is left undone and any of the two types can contribute to finishing it, the directive ones will probably say "I want you to finish it." or "I want you to help me finish it" or probably "I will do the task, not to worry." It can be seen the direct types leave no room for imagination. Informative ones are likely to say, "The task is incomplete." This statement leaves scope to be reinterpreted in various ways. Informative types are waiting for a conclusion, they just relayed information and are to see how things play into the future scenarios. So sometimes direct people can be perceived as bossy as it seems like they are ordering around. One thing that needs to be kept in mind that all archetypes can be directive and informative. But some types are primarily directive and secondarily informative and vice versa. People tend to feel comfortable using their primarily focused attributes. An informative type would be directive if necessary but they prefer to be informative usually. Worst case scenario can be for these types is that directive people can be too controlling and informative people might have a victim mindset. Structure and See It Through types are directive, meanwhile Starter and Behind The Scenes types are Informative. [7] [9] [11]

3.1.2 Initiator vs Responder

Like for directive and informative, all these components focus on primary behavior. So in social situations, there are ones to initiate a conversation and there are some to respond to the conversation to carry on. It is close to the theory of introversion and extraversion where introverts prefer solitude to crowd and the extroverts are the opposites. Introverts gain their mental energy by staying alone whereas extroverts gain their mental energy by interacting with crowds. The introverts can still maintain their energy if it is a one on one conversation. Extroverts can also do that but they get drained of mental energy sometimes if it is a one on one conversation. Acting the opposite of these two types will be acting out of their comfort zone. Structure and Starter types are the initiator and See It Through and Behind The Scenes types are the responder. In Jungian Analytical Psychology, being Introverts or Extroverts have nothing to do with outgoing. It is all about initiating or responding in social situations. [7] [9] [11]

3.1.3 Control vs Movement

The ones who favor control prefer outcomes over everything else. If a project or a task is producing results then they believe there is progress. On the other hand, the movement types prefer progress over everything else. As long as there is progress no matter how trivial it may be, it is all good. Control types want order and outcomes. Movement types can not stand stagnation. Stagnation is the bane of their existence. They would want to work things in a way so that there is some forwarding. Sometimes they bring chaos to move things forward. In this way, movement types are more chaotic than the control types. Control types might freak out if things are not in order. Their purpose is to bring order to everything. If there is chaos, there cannot be progression. Structure and Behind The Scenes types are control types and Starter and See It Through types are movement types. [7] [9] [11]

The four Interaction Styles are—

- ❏ *Structure Types*
- ❏ *Starter Types*
- ❏ *See It Through Types*
- ❏ *Behind The Scenes Types*

Structure Types:

The types from this communication style (ESTJ, ESTP, ENTJ, ENFJ) are directive, initiator, control-oriented. They are also known as “In-charge” types. Because they love to take charge. They love structure and being in control. In control does not necessarily mean they are controlling types. Structure types are pretty common in the fields of administrative jobs, entrepreneurship, leadership, financial management, etc. ENTJ types excel in business and real estate agencies. ENFJ loves to do similar administrative tasks that are more altruistic in nature and people-oriented. ESTJ types are well-known presidents. ESTP types have a habit of testing the existing rules and structures and breaking and rebuilding them. These types just cannot stand chaos. They get insecure if their plan gets out of control. If things are out of order, this bugs them a lot. [3] [7] [9] [11]

Starter Types:
The types from this communication style (ESFJ, ESFP, ENTP, ENFP) are informative, initiator, movement-oriented. These types invest their fullest potential when starting a project, and as they get near the ending, they start to get drained of that energy. These people can talk for hours if needed and still do not lose their mental energy. Starter types are super speedy at doing their task because of these components mentioned previously. Movement especially. If there is no flow of progress, it bothers them tremendously. As Informative types, they can be very cryptic. There is always context to their speech and at times it gets difficult to decrypt them. [7] [9] [11] [4]

See It Through Types:
The types from this communication style (ISTJ, ISTP, INTJ, INFJ) are directive, responder, movement-oriented. These types are also known as "Finishers" or "Chart the Course" types. They are called that because they are very focused on finishing a task. But they have a very hard time starting a task, no matter how trivial the task is. But once they start, they put all of their energy into finishing the task. These types have the highest chances of being workaholics. As direct and responding types, they love to stay informed and updated. They prefer someone else would inform because their conscious primary side is not comfortable initiating. If they do not see progress happening, they have tendencies to withdraw themselves from the project and focus on something else to chart their lists. As they are responding types, too much social interaction leads them to initiate more often and that drains them. [7] [9] [11] [5]

Behind The Scenes Types:
The types from this communication style (ISFJ, ISFP, INTP, INFP) are informative, responder, control-oriented. These types are also called "Background" types because they prefer staying in the shadows. Being around crowds for too long drains them because either because they are out of their comfort zones (Si users) or their freedom is being inhibited (Se users), and as a result, they feel like they are losing their control in the environment. Like the Starter types, they also speak in various subtexts or contexts. As control types, they love to go with their own paces. The best way for them to be comfortable around someone would be to give them personal space from time to time, they need that more than any other type. [7] [9] [11] [6]

3.2 Temperaments

The idea of Temperaments came from the theory given by Plato. [7] [11] [9] Like interaction styles, there are four temperaments to type these Jungian Archetypes. Based on Dr. Berens theory, C.S. Joseph made the terms more accessible by renaming them. [12] [9] These temperaments have various components as well—Systematic, Interest-based, Affiliative, Pragmatic, Abstract, Concrete. [7] [11]

3.2.1 Systematic vs Interest-based

Systematic people are all about systems and methods. They prefer to understand the details of any system. These said systems can be of various kinds—it could be traditional social norms or scientific methods. Systematic types can be seen too much caught up in the details to the point they forget the real purpose of the thing they are so engrossed in. Systematic people dress to impress. For example, in an interview, the systematic types are more likely to dress formally without notice. But the interest-based types might not follow this for the first time. They might start to follow this after figuring out how the structure works. Interest-based types are into the inner purpose of everything. They are the ones always asking the point of all these structures. They do not necessarily understand the structural behaviors as the systematic people do. Systematic people's primary goal in life to find comfort where interest-based ones seek happiness more often. The xSxJ (Guardians) and xNTx (Intellectuals) types are systematic types and xSxP (Artisans) and xNFx (Idealists) types are interest-based types. The rough percentage of Systematic and Interest-based people are 55% and 45% respectively. For this, society favors systematic people more than the other. Both types can be meticulous but interest-based types are less into methodologies. [7] [9] [11]

3.2.2 Concrete vs Abstract

In simple terms, Sensing people are concrete and intuitive people are abstract. Concrete types are tangibility, reality, physical world, memories. Concrete people focus on the what-is. Abstract people focus on the what-ifs, the possibilities. The concrete people have either Si or Se in the hero and parent slot in their ego side. The abstract types are primarily Ni or Ne users in the same slots. The Guardian and Artisan temperaments are concrete while the Intellectuals and Idealists are abstract. It is more likely for concrete types to store memories by taking pictures, the abstract types may not always do so. The opposite may happen but that would be a secondary choice for them. If a concrete type sees an object, they will think about its tangible properties like height, length, materialistic characteristics, durability, etc. The abstract might think about what may happen at an object in different circumstances and scenarios. They would make possible scenarios in their mind and play them mentally to draw a clear picture. The Ni users abstract types will focus on their possible future while the Ne users will focus on collective futures like where a race is heading towards for example. The rough percentage of Concrete and Abstract people are 70% and 30% respectively. These two kinds have a hard time understanding each other. Also because of the drastic difference of these two kinds, the Abstract people are often misunderstood for living in their heads. [7] [9] [11]

3.2.3 Pragmatic vs Affiliative

Pragmatic people are individualistic. Affiliative people are into authority and teamwork. Pragmatic people choose the efficient way of working, what is the most optimized method of accomplishing something. They usually end up doing the task themselves because they think others are dragging them down. Affiliative people are more team-oriented, they prefer to do things the "right" way. Their objective is to bring harmony. Sometimes because they are too agreeable, they can fall into peer pressure. Pragmatic types on the other hand are too independent and sometimes may seem too arrogant. For pragmatic types, it is always about what "works". Affiliative types always ask for permission, pragmatic types, on the other hand, would do something themselves without permission and ask for forgiveness later. The Guardian and Idealists temperaments are affiliative while the Intellectuals and Artisans are pragmatic. The percentage of pragmatic and affiliative types are 45% and 55% respectively. Because of this ratio, society favors affiliative kinds more. [7] [9] [11]

The four Temperaments are—

- ❑ *Guardians, SJ*
- ❑ *Artisans, SP*
- ❑ *Intellectuals, NT*
- ❑ *Idealists, NF*

Guardians, SJ:

The types from the Guardian Temperament (xSxJ—ESTJ, ESFJ, ISTJ, ISFJ) are past-focused. They love and protect traditions. They have Si in higher slots (Hero or Parent) which explains their higher awareness of long-term memory and duty. Si is their motivational factor. These types are heavily duty-focused. They do their job with the "should" in mind, not "wanting" (less awareness of Ne or Ni). Guardians excel in military duties. Firefighters, politicians, policemen, nurses, doctors, dentists—SJ types are seen to join such professions, because these fields have a lot to do with protecting, serving, safety, and defending. As Ne users, they are aware of other people's intentions quite well (especially ESxJ). According to statistics, 40% of the world population are Guardians. The UK, the USA, etc countries are SJ-biased societies. Most of the first world society is dominated by SJ types. The traditions are highly in favor of them.

As Si users, they have to feel comfortable. As much as they love to stay in their comfort zones, they also love having new experiences because Si loves experiences and making memories. They have well-developed inner discipline thanks to which they can be very focused. Many qualified choreographers are of the SJ temperament. Guardians are systematic, affiliative, concrete. They love attending social parties and are very serious about it. They have the mindset of "you only live once". These Si users have long time memory access and high mental tolerance; they remember everything. Also, loyalty is very important to them than any other type. They can not let go and are loyal to a fault. Also, because of high endurance, they can easily fit into adverse situations.

Si users naturally live in the past and have difficulty living at present. ISxJ types have more trouble living in the moment, ESxJ types have Si in their parent slot so they have less trouble doing so. As concrete, they are very focused on "what is". They firmly believe "what has happened before will happen again. History repeats itself." This belief is based on prescience affected by the Si-Ne axis. Guardians are all about their own experience and that kind of makes them more conservative and conforming. They love routine and regularity and are risk-averse; prefer to play safe. This behavior sometimes makes them control-oriented more than others.

The reason behind the high percentage of SJ temperament from an evolutionary standpoint is that these "guardians" are there to protect social values. They are more grounded and that is an important element for the progress of any race. [7] [9] [11]

Artisans, SP:

The types from the Artisan Temperament (xSxP—ESTP, ESFP, ISTP, ISFP) are present-focused. These types are freedom-based creators. They are very creative, tactical, have high mechanical awareness (Hero or Parent Se), and can manipulate the physical environments around them well. They are good at building and fixing things, for example working with cars, running an excavator, construction, drawing, and designing, etc that need to use hands and tools. Artisans are all about living in the moment being spontaneous. They are also about what they want (Child or Inferior Ni). Personal freedom is very important to them. Artisans are pragmatic, interest-based, concrete types. They are very individualistic people. Africa is an SP-biased society. The world population has 30% of Artisans. Artisans love to challenge the existing norms. They are the master of a reality check, their high Se helps in that. Ni is their motivational factor. So these types need freedom of choice more than any other temperaments. If they are not given that and feel confined, they will get ragey about it and might explode. They feel like they always have to react at the moment because they might forget about it and lose the energy. This is due to the short-term memory caused by Se. So to them, it is always "now or never". Their approach is about the physical confrontation at the moment, with tactics. Since they are low Ni users, they do not plan ahead of their time. It can also be said, low Ni and high Se are about improvisation. Planning makes them feel trapped and that is a hindrance to their freedom of choice. This is why they are very spontaneous.

These types appreciate when they are given that choice. So children of these types might not behave well when things are imposed upon them. SP types are reality-focused, their Se and pragmatism make them more practicality oriented than theory-oriented. They need to experience something to understand the underlying patterns. The abstract type might focus on the possibilities to understand something, concrete types, especially these types need to see if it has really happened or it is just a theory. They will not believe in something unless they see it in physical form. They can be indecisive, they take a long time to make a decision, whether it is based on logic (Ti) or morality (Fi). These Se users love to give others a good experience. Se is focused on performing, showing others something (mostly creativity), so a

lot of Artisans join the entertainment industry. Lower Ni makes them prone to giving in to the moment at times.

Creating is everything to them, and their firm belief is without freedom, creativity is not possible. SP types love construction, media creation, art, crafting—these are about trying to tear down a structure to rebuild it. Their art is the way to tell the truth. They love to create the ultimate experience. A good way to deal with the Artisans is to outlast them, by allowing them to rage out at the moment. Their Se forgets things easily as new data always comes to their internal system by pushing out the older ones. SJ types are good at handling the Sp types. These two temperaments are highly socially compatible. In fact, SJ types prefer to interact with SP types more. [7] [9] [11]

Intellectuals, NT:

The types from the Intellectual Temperament (xSxP—ENTJ, ENTP, INTJ, INTP) are future-focused. Their focus regarding the future is on the systems. These types are future intuitive thinkers who use their logic (Ti) or rationale (Te) based on their types. They are very objective, theoretical, and always aware of the cutting edge. They process their thinking focusing on the possibilities to improve systems. So they are very much into the what-ifs. They love innovation. The xNTJ types (Se users) combine mechanical aptitude with intuition and innovate new things. The xNTP focuses on the possibilities of understanding where the human race is heading towards (Ti and Ne) as a species. The focal point of these possibilities could be about a certain technology or a system. NT types are the most compatible with types from their own temperament (xNTJ with xNTP).

Intellectuals are intrinsic in nature. They are systematic, pragmatic, and abstract. 15% of the world population are from this temperament. Japan is an NT-biased society. Unlike SJ (Si users) and SP (Se users) types, this temperament can not be identified based on a single function. There are either Ni+Te (xNTJ types) or Ne+Ti (xNTP types). Since they are future-oriented thinkers, they want to bring systematic changes in society by constantly testing the boundaries of what is possible. They love to challenge the boundaries and rules with their process-driven perspectives. Artisans also do that but they are more concrete in nature so their change is based on the existing reality. NT people are more into adding new flavor in reality. For example, the theory of relativity and quantum mechanics were proposed by people from this temperament. This is why the Intellectuals are good at making blueprints. Since Intellectuals are so little in number and percentage, they are often misunderstood by other types. There are several biases against these types because of their abstract thinking as it is perceived as unrealistic. [7] [9] [11]

Idealists, NF:

The types from the Idealist Temperament (xSxP—ENFJ, ENFP, INFJ, INFP) are also future-focused. But their focus regarding the future is on the values and social norms, basically people-oriented. Like NT types, NF types are also towards innovation and

improvements but that is focused on social change. People like Gandhi, Martin Luther King Jr are from this temperament. The xNFP types are always advocating ideas with Ne and Te. These types can easily and efficiently change the belief system of a society with the help of these two functions. They are also good at marketing, sales, and finance. While xNFJ types are executors (with the help of vision gained from Ni and Ti). Idealists are affiliative, interest-based, and abstract. 15% of the world population are from this temperament.

Idealists are intuitive feelers. They are high on sympathy (Fi) or empathy (Fe). They are very people-focused. So many of the NF types dedicate themselves to volunteer work. The xSFJ types do that too. But they love to protect whereas Idealists love to improve. Many ENFJ and INFP join the educational sectors. NF types are the most compatible with types from their own temperament (xNFJ with xNFP). Their motivation comes from improving society for the sake of its betterment, to make tomorrow better for the entire race. [7] [9] [11]

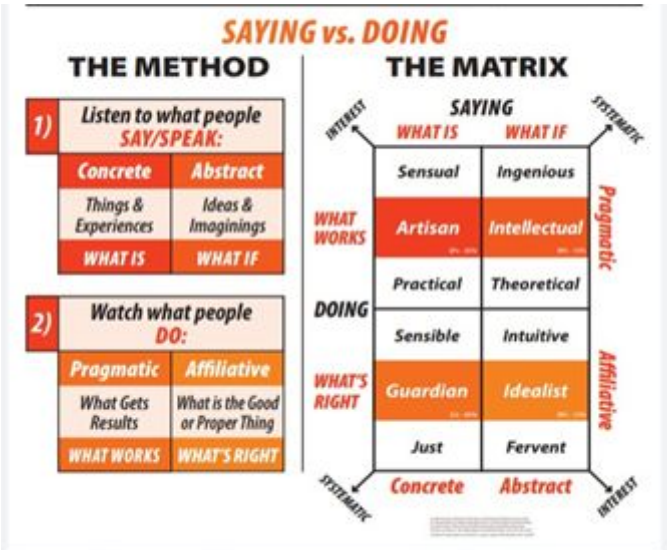


Fig. 3.1 Temperament determination matrix. [12]

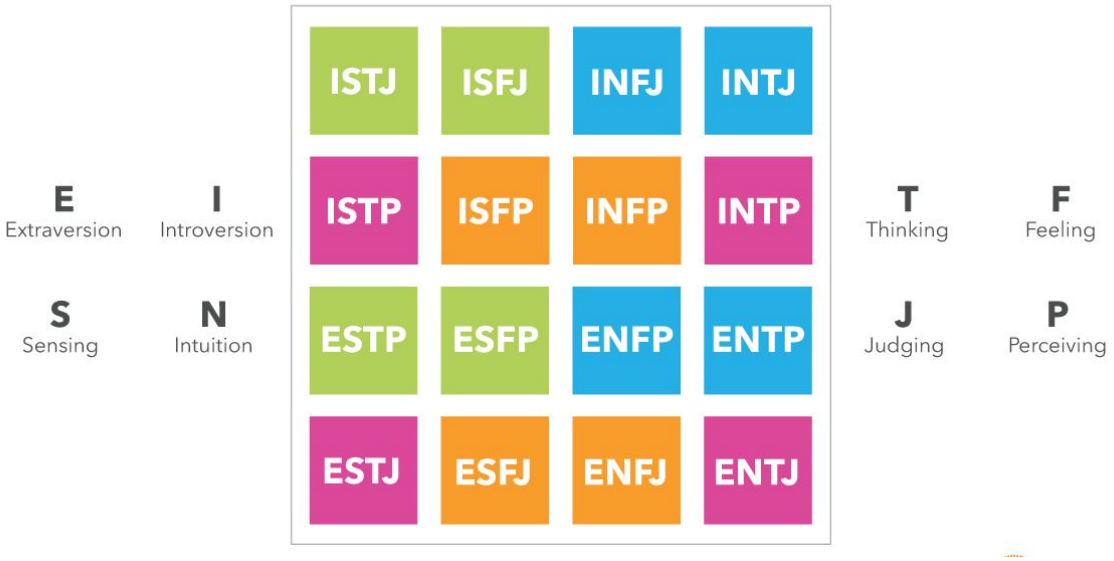


Fig. 3.2 16 personality type and temperaments. [16]

3.3 The 16 Archetypes

ESTJ (Overseer): Guardians temperament. Structure or In-charge type. So they are concrete, systematic, affiliative, and directive, initiator, control. Function order, Te-Si-Ne-Fi, Ti-Se-Ni-Fe. They are of the Philosophers Quadra. [12]

ESTP (Persuader): Artisans temperament. Interaction style, In-charge type. So they are concrete, interest-based, pragmatic, and directive, initiator, control. Function order, Se-Ti-Fe-Ni, Si-Te-Fi-Ne. They are of the Templars Quadra. [12]

ENTJ (Chief): Intellectuals temperament. Interaction style, In-charge type. So they are abstract, systematic, pragmatic, and directive, initiator, control. Function order, Te-Ni-Se-Fi, Ti-Ne-Si-Fe. They are of the Wayfarers Quadra. [12]

ENFJ (Mentor): Idealists temperament. Interaction style, In-charge type. So they are abstract, interest-based, affiliative, and directive, initiator, control. Function order, Fe-Ni-Se-Ti, Fi-Ne-Si-Te. They are of the Templars Quadra. [12]

ESFJ (Supporter): Guardians temperament. Interaction style, Starter type. So they are concrete, systematic, affiliative, and informative, initiator, movement. Function order, Fe-Si-Ne-Ti, Fi-Se-Ni-Te. They are of the Crusaders Quadra. [12]

ESFP (Entertainer): Artisans temperament. Interaction style, Starter type. So they are concrete, interest-based, pragmatic, and informative, initiator, movement. Function order, Se-Fi-Te-Ni, Si-Fe-Ti-Ne. They are of the Wayfarers Quadra. [12]

ENTP (Visionary): Intellectuals temperament. Interaction style, Starter type. So they are abstract, systematic, pragmatic, and informative, initiator, movement. Function order, Ne-Ti-Fe-Si, Ni-Te-Fi-Se. They are of the Crusaders Quadra. [12]

ENFP (Advocate): Idealists temperament. Interaction style, Starter type. So they are abstract, interest-based, affiliative, and informative, initiator, movement. Function order, Ne-Fi-Te-Si, Ni-Fe-Ti-Se. They are of the Philosophers Quadra. [12]

ISTJ (Examiner): Guardians temperament. Interaction style, See It Through type. So they are concrete, systematic, affiliative, and directive, responder, movement. Function order, Si-Te-Fi-Ne, Se-Ti-Fe-Ni. They are of the Philosophers Quadra. [12]

ISTP (Craftsman): Artisans temperament. Interaction style, See It Through type. So they are concrete, interest-based, pragmatic, and directive, responder, movement. Function order, Ti-Se-Ni-Fe, Te-Si-Ne-Fi. They are of the Templars Quadra. They have good mechanical aptitude. [12]

INTJ (Strategist): Intellectuals temperament. Interaction style, See It Through type. So they are abstract, systematic, pragmatic, and directive, responder, movement. Function order, Ni-Te-Fi-Se, Ne-Ti-Fe-Si. They are of the Wayfarers Quadra. [12]

INFJ (Sage): Idealists temperament. Interaction style, See It Through type. So they are abstract, interest-based, affiliative, and directive, responder, movement. Function order, Ni-Fe-Ti-Se, Ne-Fi-Te-Si. They are of the Templars Quadra. [12]

ISFJ (Defender): Guardians temperament. Interaction style, Behind The Scenes type. So they are concrete, systematic, affiliative, and informative, responder, control. Function order, Si-Fe-Ti-Ne, Se-Fi-Te-Ni. They are of the Crusaders Quadra. [12]

ISFP (Artist): Artisans temperament. Interaction style, Behind The Scenes type. So they are concrete, interest-based, pragmatic, and informative, responder, control. Function order, Fi-Se-Ni-Te, Fe-Si-Ne-Ti. They are of the Wayfarers Quadra. [12]

INTP (Engineer): Intellectuals temperament. Interaction style, Behind The Scenes type. So they are abstract, systematic, pragmatic, and informative, responder, control. Function order, Ti-Ne-Si-Fe, Te-Ni-Se-Fi. They are of the Crusaders Quadra. [12]

INFP (Dreamer): Idealists temperament. Interaction style, Behind The Scenes type. So they are abstract, interest-based, affiliative, and informative, responder, control. Function order, Fi-Ne-Si-Te, Fe-Ni-Se-Ti They are of the Philosophers Quadra. [12]

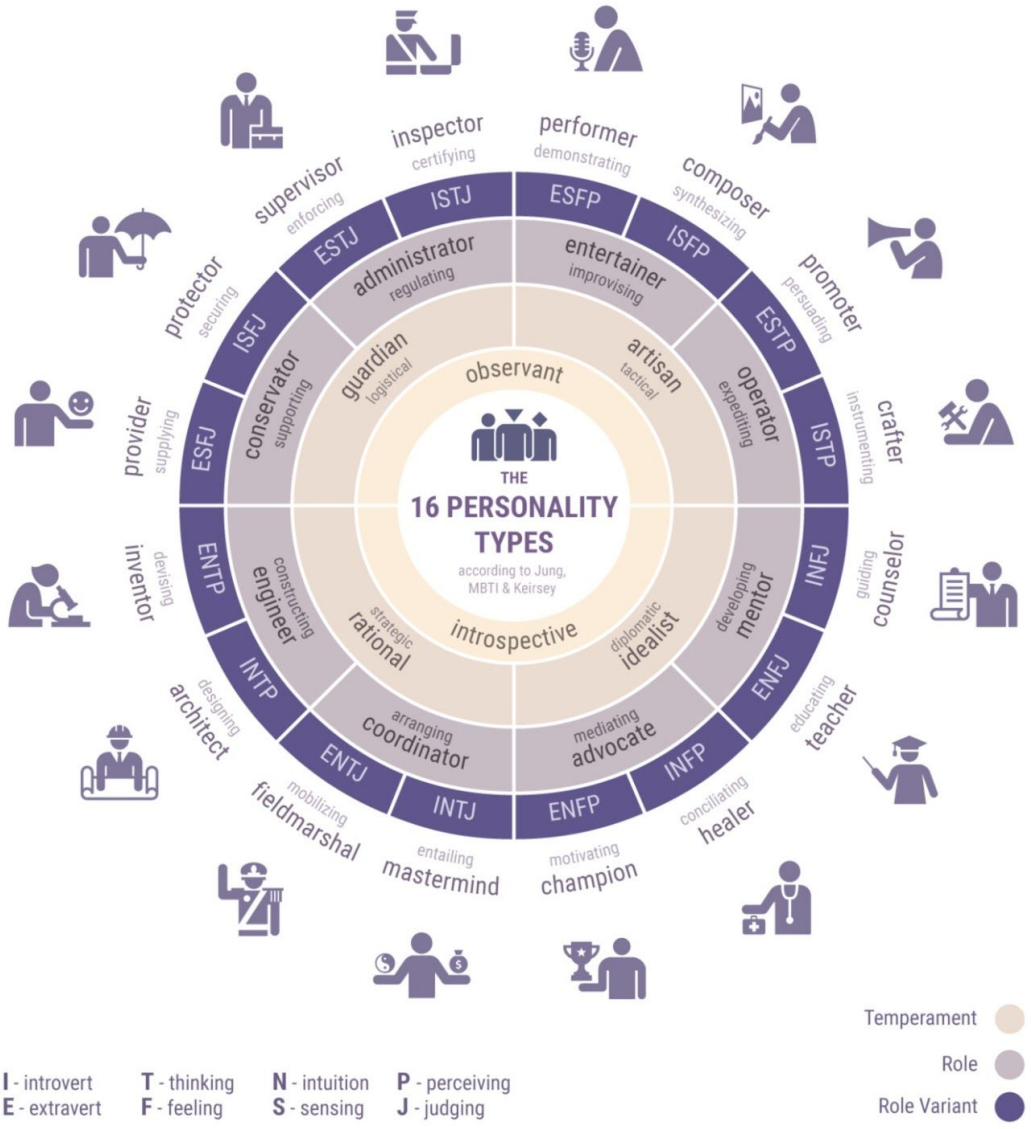


Fig. 3.3 The 16 personality types. [15]

Chapter 4

Methodology

4.1 Data Set Formation

4.1.1 Psychology survey questionnaire construction

Psychology survey questionnaires [19] are questions interrogated to gather information about an individual to assess and evaluate the mental state of the particular responder. These questions enable the scopes for a researcher to evaluate the mental traits and categorize the behaviors likewise. The studies are utilized by various establishments like medical care, corporate recruitment, enrollment firms, safeguard administrations, and some more. Such brain science review questions are made and dissected by therapists, emotional well-being experts, specialists, individuals from the legal framework, and other psychology specialists. The construction of the questionnaire for the study had the aim of executing the interrogation of individuals creating a data frame of data entries more than a thousand. For a particular and potentially utilitarian cause, the session of the questionnaire had to be taken under a revised environment and a considerable amount of time[20]. The questions are thoroughly set on three major humane psyche-based components which are the ego, the personal unconscious, and the collective unconscious. A total of 56 analytical questions were set to dissect the possibilities of detecting the psyche archetypes. The set of questions had been enacted on 73 individuals in a separated room so that the concentration of both interviewer and the responder are focused on the purpose and aim of the study.

4.1.2 Logic formulation

Logic programming [21]through Python is a programming paradigm that considers computation as automatic reasoning over a database of knowledge where information is factual and overruled through particular protocols. It is a method of programming and depends on the conventional rationale. The logic set for the assessment was solely based on the affirmative and negative responses of the individuals and conditionally dependent on particular terms those unearthed notable traits and signified archetype characteristics. Python language is highly efficient when it comes to effectively circulating a program. The foundation of the logic was established on 24 variables where all the variables were not equally functionable. Variable counts were initially null valued and increased with a positive

one or decreased with a negative one according to the stated conditions. The logic design through the Python language has the following set-up (partial functionable syntax has been mentioned):

```
a_count=0
q1_answer=input("Do you have the idea of what you want in your
life?")

if q1_answer.lower() == 'y' : a_count+=1
elif q1_answer.lower() == 'n' : a_count-=1
```

The rest of the questions follow a similar thread of logic design. All the other parameters are categorized according to the logic they have been set with. The variables that were taken correspond with the alphabet ‘a’ to ‘x’ where each of them represents a particular trait for the estimations and results.

4.1.3 Parameter determination model

The concentric components of Jungian analysis depend on multiple personality traits for example if a person is inclined towards affiliative or systematic orders or if the introspection of a person deviates with pragmatic or concrete views. The following features were used to build the data frame according to the responses of the individuals and were characterized uniquely: Ni, Ne, Si, Se, Ti, Te, Fi, Fe, Informative, Direct, Control, Movement, Initiate, Respond, Structure, Starter, See it Through, Behind the scenes, Abstract, Concrete, Pragmatic, Affiliative, Systematic and Abstract. The mentioned features widely impact on which particular quadra an individual might have their traits included in.

4.1.4 Design of dataset

The design had been initiated to categorize all the features and have the archetype well determined so that the machine learning could analyze the overall data frame and determine a recognizable accuracy. Microsoft Excel is the data visualization and analysis tool we used to prepare the entries. The responders individually have been designated to integer values that represent the features. The 16 Archetypes correspond to the integers 1 to 16 respectively. One particular Archetype ‘ENTP’ was not found in the survey. Hence data scarcity affected the overall result and the design did not include the 7th type ‘ENTP’.

id	age	Ni	Ne	Si	Se
1	23	5	9	6	16
2	22	2	4	5	11
3	22	6	2	6	2
4	22	5	9	7	7
5	23	3	9	5	7
6	24	3	7	6	12
7	23	6	3	4	8
8	23	3	8	10	14

Table 4.1. A part of the dataset in Microsoft Excel.

4.2 Machine learning techniques

4.2.1 Introduction

Machine learning is the study of getting computers to act without being unequivocally modified and programmed [22]. It is a method of data analysis that automates analytical model building and is seen as a part of AI. In the previous decade, AI has given us self-driving vehicles, reasonable discourse recognition, viable web search, and an inconceivably improved comprehension of the human genome. Machine learning is so unavoidable today that you most likely use it many times each day without knowing it. Numerous specialists additionally think it is the most ideal approach to gain ground towards human-level artificial intelligence. The study includes automatic procedures and gains work on actualizing them and getting them to work for practical purposes. All the more significantly, the hypothetical underpinnings of learning aren't confined in executions of computers it additionally opens the path to acquire the functional skill expected to rapidly and capably be applicable for strategies to practical and new issues. The techniques are Silicon Valley's highly accepted practices in development and innovation as it pertains to AI and computer-based intelligence.

Machine learning techniques are aimed at teaching machines to perform a cognitive activity, similar to the human mind using particular types and formatted data sets[23]. Machine learning algorithm, likewise called model, is a numerical articulation that addresses information with regards to a problem, regularly a business issue. The point is to go from information to understanding. For instance, if an online retailer needs to expect deals for the following quarter, they may utilize an ML algorithm that predicts those deals dependent on past deals and other important information. Also, a windmill maker may outwardly screen

significant gear and feed the video information through calculations prepared to distinguish risky breaks. For the study, we focused on using ML techniques on Neuro-Linguistic Programming (NLP). The techniques are categorized mainly in two particular classifications: Supervised Learning and Unsupervised Learning. The classification follows the mentioned flowchart.

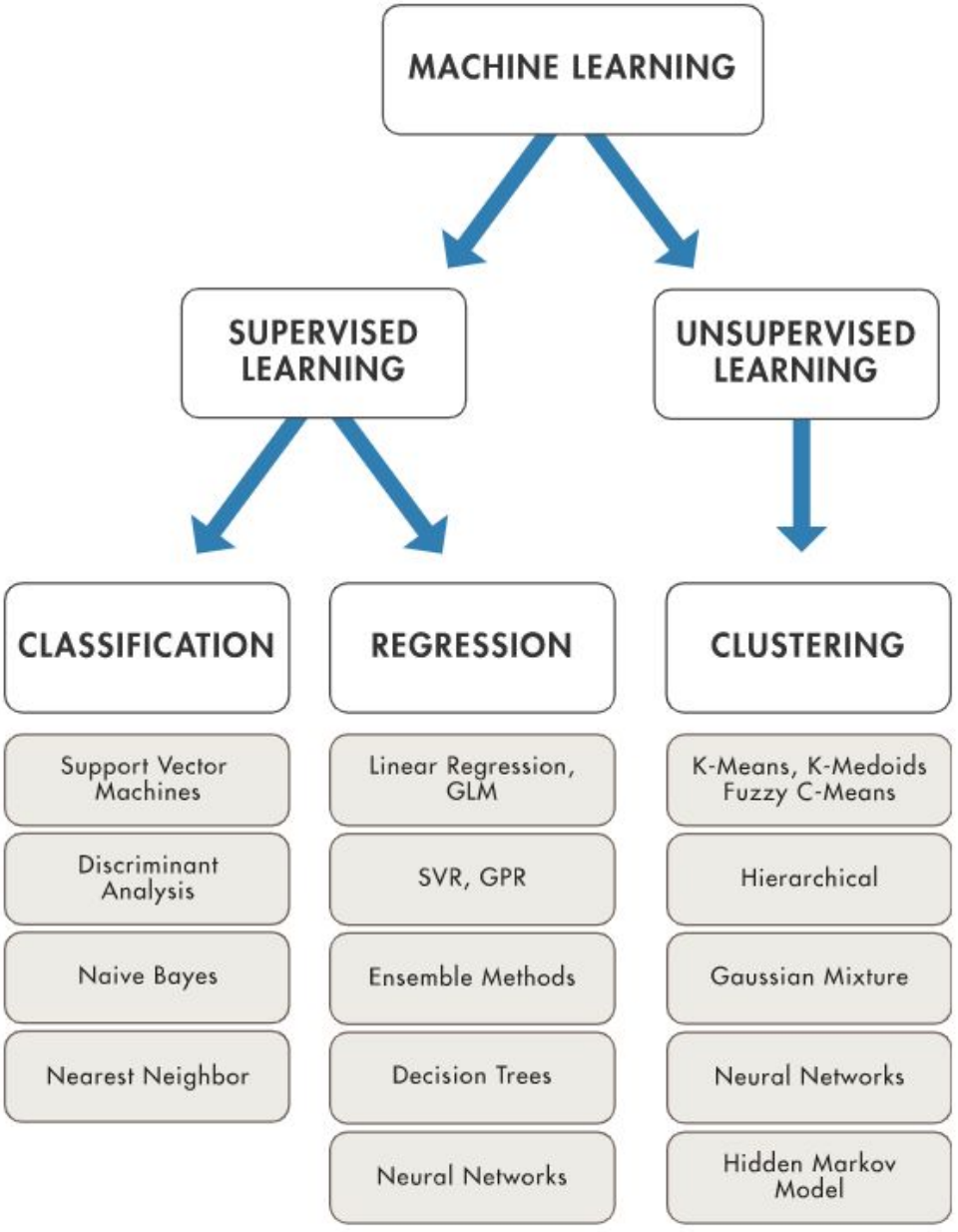


Fig. 4.1 Flowchart of machine learning and its classification.[24]

4.2.2 NLP

Neuro-Linguistic Programming (NLP) is an assortment of procedures for character improvement. Meta programs, which are ongoing methods of contributing, arranging, and separating the data found in our general surroundings, are an imperative factor in NLP[25]. Contrasts in meta programs bring about huge contrasts in conduct starting with one individual then onto the next. Character and personality types can be perceived through using and investigating meta programs. There are various techniques to anticipate character types dependent on meta programs. The Myers–Briggs Type Indicator (MBTI) is right now considered as quite possibly the most mainstream and solid technique. In this study, another machine learning strategy has been produced for character type forecast dependent on Jungian analysis, an advanced format of NLP. The presentation of the new procedure introduced in this study has been contrasted with other existing strategies.

4.2.3 Supervised learning

The major sector of practical machine learning utilizes supervised learning[24]. Supervised learning is the technique where there are input variables (x) and a yield or output variable (Y) and then is utilized with an algorithm to learn the mapping function from the input contribution to the yield of the output.

$$Y = f(x)$$

The objective is to surmise the mapping function to such an optimization point where for a particular input data (x) the output data (Y) for that information are easily anticipated.

It is called supervised learning in light of the fact that the process of an algorithm learning from the prepared and training dataset can be considered as an instructor overseeing the learning interaction. The correct answers are known, the algorithm further iteratively makes forecasts on the training data and is revised by the instructor. Learning stops when the learning accomplishes a worthy degree of execution.

Supervised learning issues can be additionally assembled into regression and classification problems.

The classification problem is the point at which the output variable is a category, for example, "black" "white" or "depressed" "not depressed".
Regression problem is the point at which the output variable is a genuine value, for example, "mass" or "capital".

Some normal kinds of issues based on top of classification and regression incorporate recommendation and time series prediction separately. There are few mainstream instances of supervised ML algorithms. Linear regression is mostly used for regression problems. Random forest is a classic choice for classification and regression problems. Support vector machines are an alternative and effective function for classification problems.

Random forest classifier

Classification methods predict or clarify a class value[26]. For instance, they can help foresee whether an online client will purchase an item. The yield can be yes or no: purchaser or not purchaser. In any case, the techniques aren't restricted to two classes. For example, a classification technique could assist with evaluating whether a given picture contains a vehicle or a truck. For this situation, the output will be 3 distinct qualities: 1) the picture contains a vehicle, 2) the picture contains a truck, or 3) the picture contains neither a vehicle nor a truck.

The most sophisticated classification algorithm is logistic regression — which makes it seem like a regression strategy, however, it's definitely not. Logistic regression appraises the likelihood of an event of an occasion dependent on at least one data source aka input.

For example, logistic regression can take as data sources two test scores for an understudy to gauge the likelihood that the understudy will get conceded to a specific school. Since the estimation is a probability, the output is a number somewhere in the range of 0 and 1, where 1 addresses total conviction. For the understudy, assuming the assessed likelihood is more prominent than 0.5, we foresee that the person will be conceded. In the event that the assessed probability is under 0.5, we foresee the admission of the individual will be declined.

Random forest, similar to its name infers, comprises an enormous number of individual decision trees that work as an ensemble. For random forest classifier each assorted tree gives out a class prediction and the class which has the most sensible votes results in model prediction.

The central idea driving random forest is a basic yet intriguing one — the astuteness of the crowds. In accordance with what data science refers to is said to be successful effectiveness

through random forest models when it is executed. Countless generally uncorrelated models (trees) working as an advisory group will beat any of the individual constituent models.

The key is the low correlation among the models. Very much like how ventures with low connections (like stocks and securities) meet up to shape a portfolio that is more prominent than the amount of its parts, uncorrelated models can deliver ensemble forecasts that are more precise than any of the individual predictions. The purpose behind this magnificent impact is that the trees shield each other from their individual errors (as long as they don't continually all blunder a similar way). While a few trees might not be right, numerous different trees will be correct, so as an ensemble the trees can move the right way. There are few prerequisites for random forest to work efficiently and well. There should be some genuine signal in our features so that models constructed utilizing those features show improvement over irregular speculating. Low correlation is expected from the predictions made by the individual trees among themselves.

4.2.4 Unsupervised learning

Unsupervised learning executes with input data but no corresponding output variables[24]. The objective of unsupervised learning is to demonstrate the hidden design or appropriation in the information to study the data. Unsupervised learning isn't based on any correct answer or any instructor. Algorithms are required to execute the structure of the data and its result on their own.

Unsupervised learning issues can be additionally gathered into clustering and association issues.

Clustering problems define the place where there are innate groupings in the information, like gathering clients by buying conduct.

Association problem learning issue is the place where it is decided to depict huge parts of your information, for example, individuals that purchase X likewise will in general purchase Y. For unsupervised learning algorithms, few things need to be true. k-means clustering should be implemented for grouping issues. Apriori algorithm needs to be used for association rule learning issues.

K-means clustering

Clustering methods direct towards the classification of unsupervised ML on the ground that they are likely to cluster or make a group perception that has comparable qualities. These

methods don't utilize output data for preparing, rather let the calculation characterize the output. The strategies of clustering methods include the utilization of visualization to examine the nature of the solution[27].

The most famous clustering technique is K-Means, where "K" addresses the quantity of clusters that the client decides to make. There are different strategies for picking the estimation of K, for example, the elbow strategy. The tactics followed by K-Means which manages the data processing focuses on few particulars. K centers are generally chosen in a random sorting method within the data. This method allocates every data point to the nearest of the arbitrarily originated centers. Additionally, it re-processes the focal point of each cluster.

The cycle finishes when the centers or focuses don't change or change on a minimalist scale. A protocol is maintained to set an iteration number with maximum value to prevent an infinite loop if centers keep on changing. If the state deviates then each data point is assigned once again.

Description:

A set of observations is (x_1, x_2, ..., x_n), every observation is a *d*-dimensional real vector, *k*-means clustering aims to partition the *n* observations into *k* ($\leq n$) sets S = {S_1, S_2, ..., S_k} so as to minimize the within-cluster sum of squares (WCSS)[27]. The goal is to find the following argument:

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x}\in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2 = \arg\min_{\mathbf{S}} \sum_{i=1}^{k} |S_i| \operatorname{Var} S_i$$

where $\boldsymbol{\mu}_i$ is the mean of points in S_i. The above-mentioned equation refers to the equivalence of minimizing the pairwise squared deviations of points in the same cluster:

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \frac{1}{2|S_i|} \sum_{\mathbf{x},\mathbf{y}\in S_i} \|\mathbf{x} - \mathbf{y}\|^2$$

The identity from which the equivalence was deduced is:

$$\sum_{\mathbf{x}\in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2 = \sum_{\mathbf{x}\neq\mathbf{y}\in S_i} (\mathbf{x} - \boldsymbol{\mu}_i)^T (\boldsymbol{\mu}_i - \mathbf{y})$$

The reason the identity makes a full impact is that the total variance is constant and it is analogous to maximizing the sum of squared deviations between points in different clusters and follows the law of total variance. This is the between cluster sum of squares (BCSS).

Algorithm:

The algorithm that has been used mostly is an iterative refinement technique. Another name for it is "the *k*-means algorithm" and it characterizes the omnipresence of the algorithm. In the hub of computer scientists, it is also alluded to Lloyd's algorithm and naive k-means for its obscurity and fast-found alternatives[27].
For an initial set of *k* means $m_1^{(1)},...,m_k^{(1)}$, the algorithm proceeds in two steps through alteration:

Assignment step:

This step is to assign each observation to the cluster with the nearest mean: that with the least squared Euclidean distance[27].

$$S_i^{(t)} = \left\{ x_p : \left\| x_p - m_i^{(t)} \right\|^2 \leq \left\| x_p - m_j^{(t)} \right\|^2 \ \forall j, 1 \leq j \leq k \right\}$$

Here a particular x_p is directed towards only a particular $S^{(t)}$ when it could be assigned to many more of them.

Update step:

In this step, the recalculation of means or centroids is observed being assigned to one particular cluster[27].

$$\sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2 = \sum_{\mathbf{x} \neq \mathbf{y} \in S_i} (\mathbf{x} - \boldsymbol{\mu}_i)^T (\boldsymbol{\mu}_i - \mathbf{y})$$

The calculation has been combined when the tasks presently don't change. The calculation isn't ensured to track down the ideal.

The algorithm is regularly introduced as appointing objects to the closest cluster by distance. Utilizing an alternate distance function other than Euclidean distance will optimize the possibilities of the algorithm not converging. Different adjustments of k-means, spherical k-means, and k-medoids have been proposed to permit utilizing other distance measures.

Initialization Methods:

Few particular initialization methods are found to be utilized for k means algorithm. Mostly used strategies are Forgy and Arbitrary Partition[27]. The Forgy strategy haphazardly picks k observations from the dataset and utilizes these as the underlying means. The Random Partition technique first haphazardly appoints a group to every observation and afterward continues to the update step, in this way registering the underlying intent to be the centroid of the bunch's arbitrarily doled out focuses. The Forgy technique will in general spread the underlying methods out, while RP puts every one of them near the focal point of the data index. As per Hamerly et al., the RP strategy is for the most part ideal for algorithms like the k-harmonics means and fuzzy k-means. The Forgy strategy for introduction is ideal when it works for standard cluster algorithms and expectation maximization.

4.3 Execution of k means clustering algorithm

4.3.1 Platforms and libraries

The platform that has been used to execute the desired program was Google Colab. Particular libraries were imported. Scikit-learn machine library was used to functionalize the entire clustering method. The scikit-learn is also known as sklearn is the most efficient machine learning programming library for python. This library is highly occupied with most of the utilized functions for program execution and result management. The KMeans algorithm was later on imported from this particular library. The other libraries that have been used are pandas, matplotlib.pyplot, StandardScaler from sklearn.preprocessing, and Axes3D from mpl_toolkits.mplot3d.

4.3.2 Data framing

The data set that has been originated through logic algorithms and implemented on Microsoft Excel is converted into .csv file and imported and then read through pandas program. The data frame has been created by composing discernable information blocks that have been created and likewise prepared for further optimization. The structure of the data frame in the model is 72 rows × 25 columns. A short demonstration of which is given as follows:

	id	age	Ni	Ne	Si	Se	Ti	Te	Fi	Fe	Informative	Directive
0	1	23.0	5	9	6	16	3	6	3	2	0	1
1	2	22.0	2	4	5	11	3	0	3	3	0	1
2	3	22.0	6	2	6	2	3	3	2	2	1	0
3	4	22.0	5	9	7	7	5	3	1	4	1	0
4	5	23.0	3	9	5	7	4	4	1	4	0	1
...	...	...	...	...	...	...	...	...	...	...	...	...
67	68	NaN	10	5	13	7	3	0	1	4	1	0

Table 4.2. Optimized data frame using pandas in Google Colab.

For data framing, the character and convolutional-based symbols didn't have the permission to be executed while data analysis.

4.3.3 Data cleaning

Data cleaning refers to the process where unnecessary or insignificant data are detected, corrected, removed, or controlled. For data cleaning first a check and run was done to detect if there were any null values present in the data frame. The feature ‘age’ according to Jungian analysis imposes very minimal correlation. Upon null checking, it was found that the feature ‘age’ contained most of the null data entries. So the detected feature was then dropped so that it didn’t have any impact on the cluster formation

4.3.4 Data preprocessing

The data preprocessing subsequently consisted of two vital steps namely data integration and data transformation. The data set which was imported was later on integrated as a singular module for further processing. The Standard scaler algorithm used from sklearn library was used to transform the data set into sets of arrays to complete the entire data mining.
Cluster number securing:

One of the major goals of this model was to detect 16 distinguishable clusters for 16 types of Archetypes. Since our data set lacked one entire Archetype classification The cluster number was hence then was set to 15 to get 15 authentic clusters through the processes the data frame has to go through. The following syntax was used to define the cluster number from the preprocessed array:

```
clusterNum = 15
k_means = KMeans(init = "k-means++", n_clusters = clusterNum,
n_init = 64)
k_means.fit(X)
labels = k_means.labels_
print(labels)
```

A random n initial value of 64 was taken to check the Euclidean distance of the data points to set distinct clusters. The labeled points were cross-checked and defined in a particular label string to, later on, generate clusters and necessary suitable plots.

4.3.5 Graphs and plots

Scatter plot:

A scatter plot was generated by importing NumPy library to visualize the relation between the features 'Archetype' and 'id' where id refers to all the particular entries of the data set.

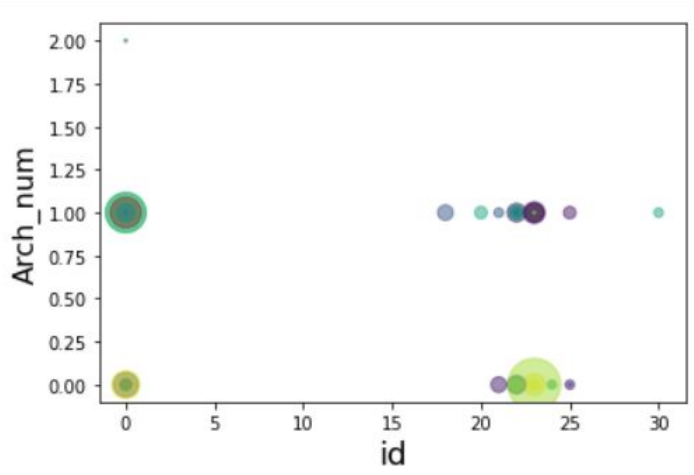


Fig. 4.2 Scatter plot of Archetype Vs ID

The plot here represents the 15 different archetype densities according to the ids of which the survey was taken.

3D plot:

The 3D cluster plot was set while emphasizing on the clusters to be formed. The desired figure was aspired to have 15 different clusters. But the data optimization needed at least

1000 entries to have distinguishable clusters. Following are two 3D plots simulated from the data entries from the set:

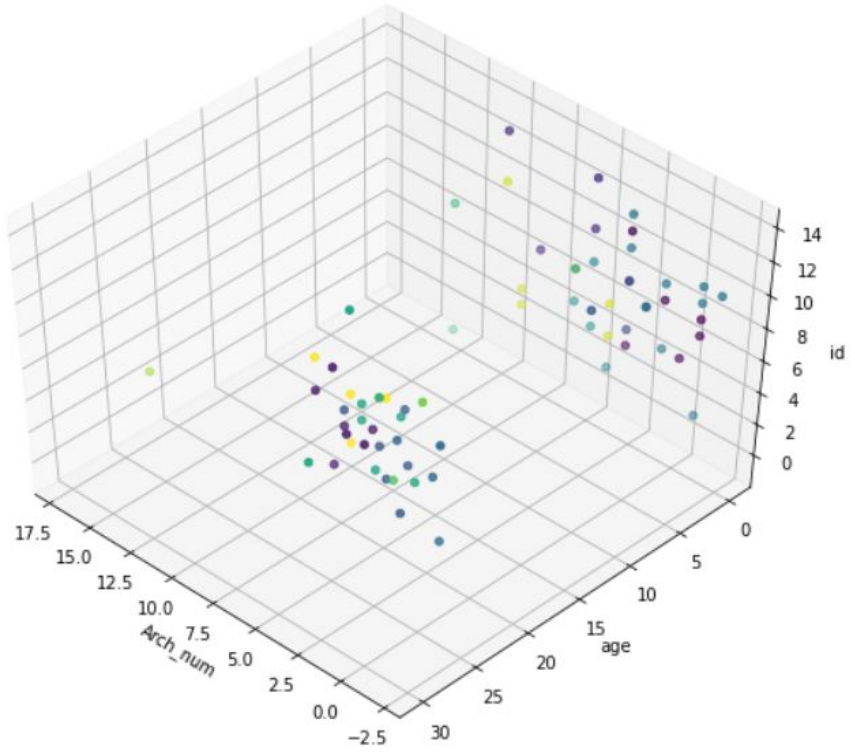


Fig. 4.3 3D cluster plot (Arch_num, age, id)

The first 3D plot was done giving priority to Archetype number with respect to age and id as initially when the data recovery was done the null values were replaced with necessary integer values. Later on, the data significance was implemented and age was found not to have an impact on the result hence the most significant nature was evaluated with the 3D plots. The following shows the relation of Archetype, the trait ‘Ne’ and id, and the minimal scattered clusters are found likewise.

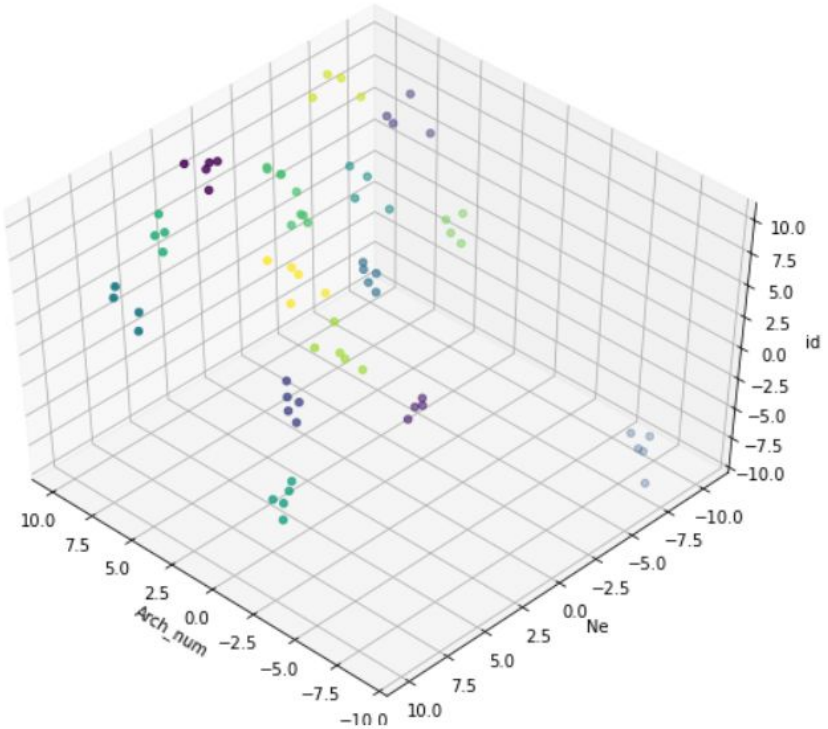


Fig. 4.4 3D cluster plot (Arch_num, Ne and id)

The different data points and colored dots represent different Archetypes and the clusters formed. This plot is a report of better data sort and cluster formation than the previous plot results.

Blob graph:

The blob graphs which are generated through Gaussian distribution helps bring out the characteristics of the features that have been processed throughout the model. The following figures are blob plots generated from the feature and defined clusters:

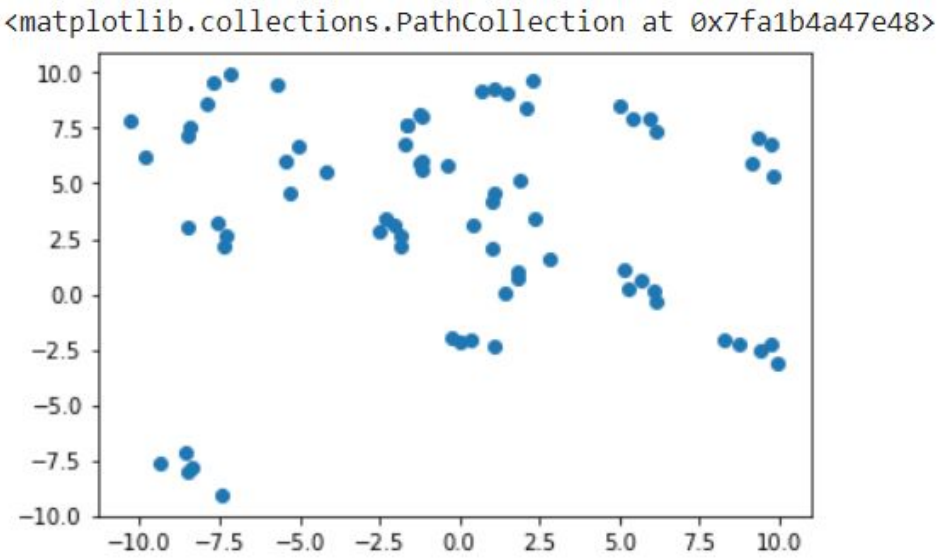
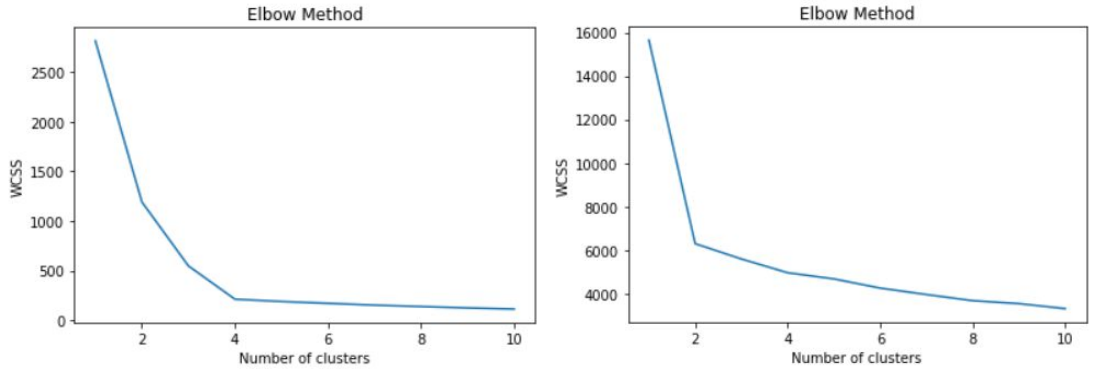


Fig. 4.5 2D Blob clustering plot of Arch_num.

Elbow Method and WCSS:

The significance of the elbow method lies in implementing k means clustering on a particular data set for a range of values of k. This method is basically used to determine the number of clusters in the procession. For each value of k an average number is assorted for each cluster. The strategy consists of plotting the clarified variation as a component of the quantity of clusters and picking the elbow of the curve as the quantity of the clusters to utilize. The connection between the quantity of clusters and Within Cluster Sum of Squares (WCSS) are charted at that point where the quantity of clusters is selected. The change in WCSS starts to level off (elbow method) at that particular phenomenon. WCSS is characterized as the amount of the squared distance between every individual from the cluster and its centroid.



(a) (b)

Fig. 4.6 Elbow Method Vs WCSS (a) for k=1000 (b) for k=750.

Initially, the value of 'k' was randomly chosen and the max iteration number was kept to 1000 and 750 for the respective figures mentioned above. Afterward, the iteration number was kept at 1500 and the value of k was kept at 15 and the following figure was generated.

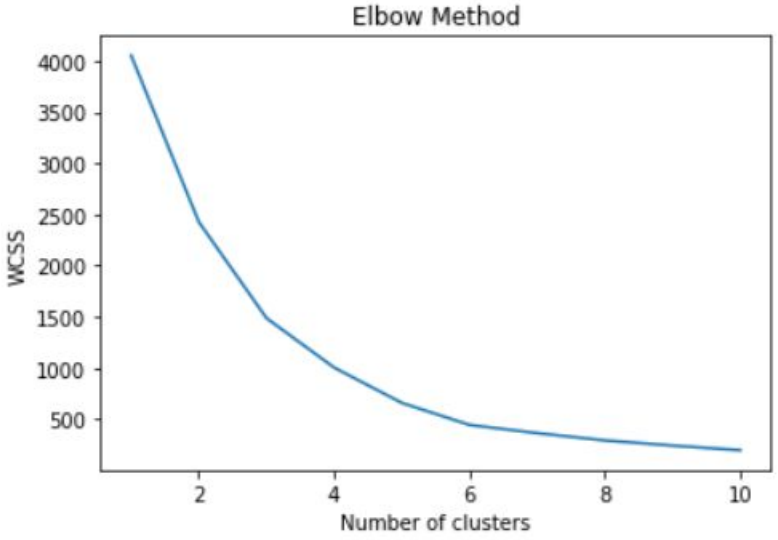


Fig. 4.7 Elbow Method Vs WCSS for k=1500

4.4 Implementation of Random Forest Model

4.4.1 Platforms and libraries

The libraries that were used to run the supervised model are pandas, NumPy, Sklearn, matplotlib.pyplot, re, seaborn, plotly, warnings. For visualization and scoring SVC, StandardScaler and Kfold were imported from Sklearn and Axes3D from mpl_toolkits. The configurations from Sklearn Metrics were further used to determine prediction, feature significance, and accuracy of the model.

4.4.2 Data frame distribution (test and train)

For this particular model, the entire dataset was divided into two parts. As it has been discussed previously in the analogy that for supervised models and algorithms, an instructor is needed. That's what's done when it comes to verifying the data set and its module for a further procession. The two parts are test and train. The train set .csv file reads the ids and their Archetypes whereas the Archetypes are absent for the test set. The train set contains the

exact number of data the test set contains for data scarcity. The following plot shows the number and type of Archetype in the survey:

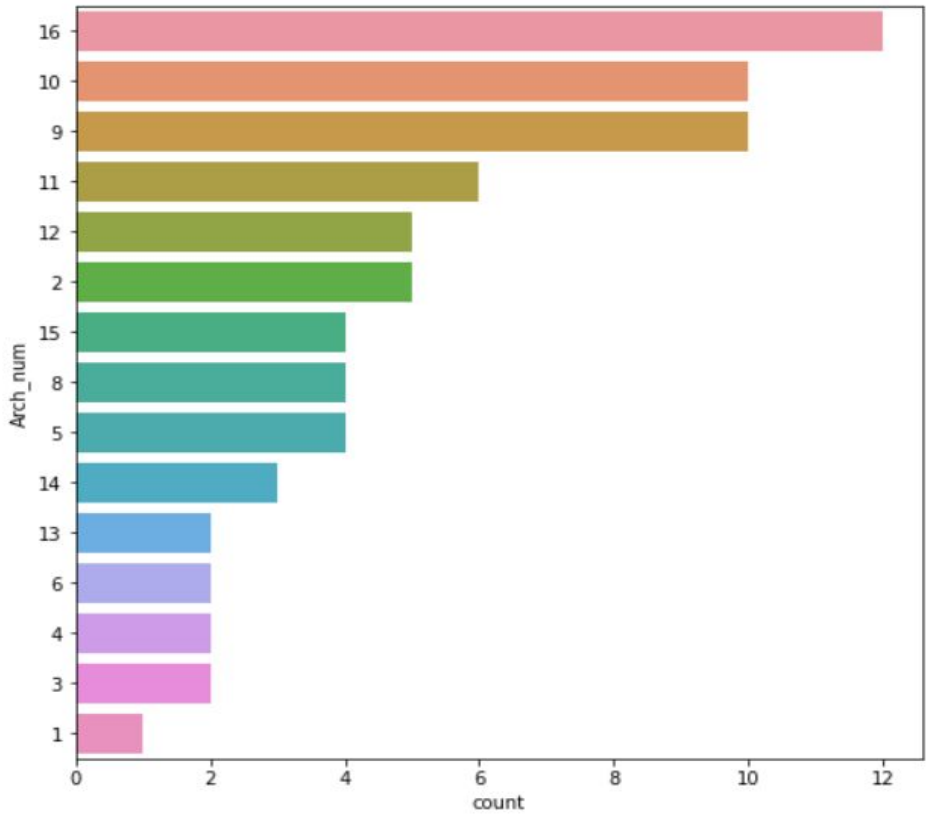


Fig. 4.8 Archetype count for the generated dataset (Arch_num count).

4.4.3 Data cleaning

The null values were checked for the entire dataset to be pre-processed. The column age once again lacked information for the test set. But for this particular model the feature age had its null values replaced with suitable integer values to keep the inter feature correlation at its minimum.

4.4.4 Setting unique feature

The features ‘id’ and ‘Arch_num’ were individually unique for different entries. Hence these two valid features were designated as unique characters or traits for further execution of training. The train set was then composed of identity characters for id and Arch_num. The feature ‘id’ was unique and had a low side impact on accuracy gathering of similar sorted characters; it was dropped for particular program execution.

4.4.5 k-NN imputation

The kNN imputation method ensures a characterized and categorical absent value to be ascribed with the greater part among its k nearest neighbors, and the normal worth (mean) of the k nearest neighbors is viewed as the expectation for a mathematical missing value, called as mean rule. The knn imputation was therefore executed on the dataset for further distinguishable values in the array.

4.4.6 Pearson Correlation Mapping

This mapping visualization was solely based on the Pearson correlation coefficient that follows the following algorithm where it is a measure of linear correlation between two sets of data. Here,

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

r = *correlation coefficient*
x_i = *values of the x − variable in a sample*
$\bar{x}$ = *mean of the values of the x − variable*
y_i = *values of the y − variable in a sample*
$\bar{y}$ = *mean of the values of the y − variable*

The correlation forms the visual representation of the features having a particular correlation coefficient with other features.

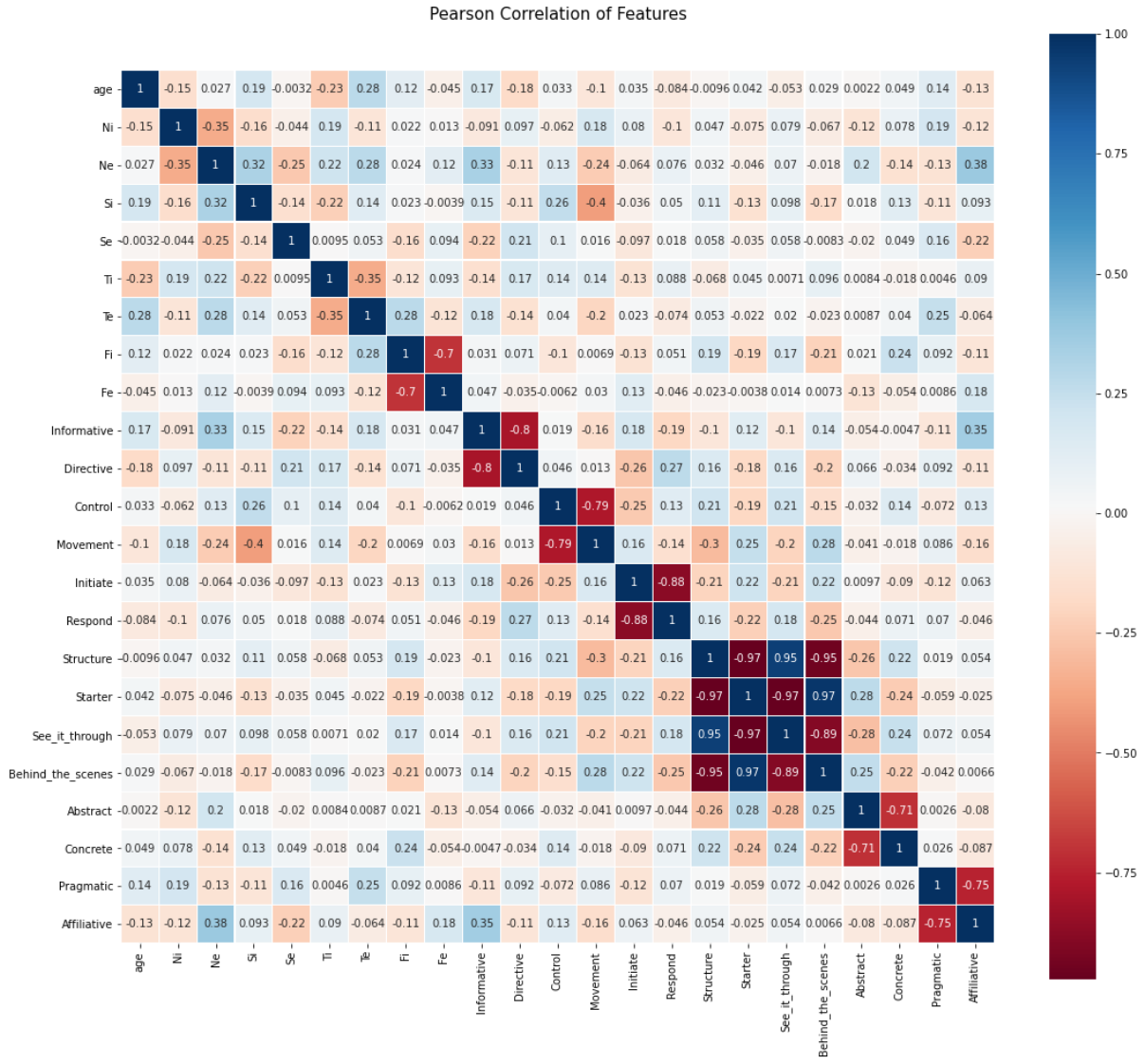


Fig. 4.9 Pearson Correlation Map (for 23 features)

4.4.7 Random Forest Classifier

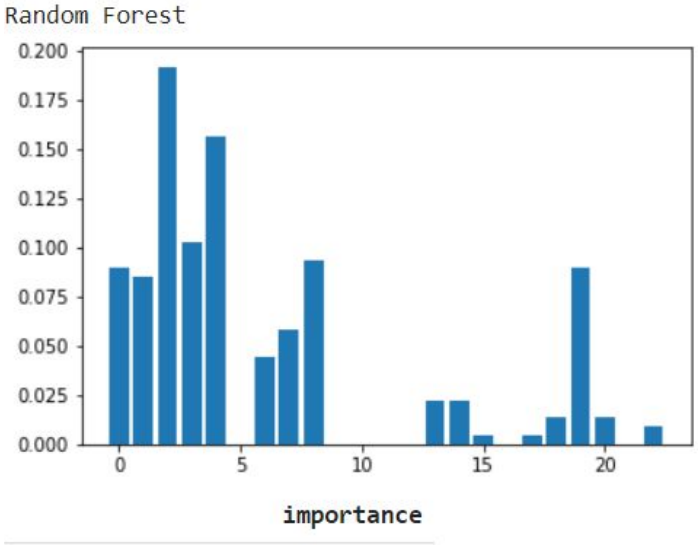
The classifier ensures the relatable operations that will run the entire train set for the test set to give a satisfactory accuracy. The characters that were defined in the program for the classifier is mentioned as follows:

```
seed=100;
rf = RandomForestClassifier(
                        min_samples_leaf=20,
                        n_estimators=500,
```

```
bootstrap=True,
oob_score=True,
n_jobs=-1,
random_state=seed,
max_features='auto')
```

4.4.8 Feature Selection and Importance

This step of the model deals with the determination and calculation of the significance of a definite feature while randomizing it with a particular value. The weight of the values given to definite features would define a satisfactory competency if the scores were near `~0.9, the highest feature importance that was found in the study was for 'Ne' with a feature importance score of 0.1912 which is a very low value and was found due to data scarcity



Ne	0.191964
Se	0.156250
Si	0.102679
Fe	0.093750
age	0.089286
Abstract	0.089286
Ni	0.084821
Fi	0.058036
Te	0.044643
Initiate	0.022321

Fig. 4.10 Feature importance and validation value.

4.4.9 RF score

RF-Score is a first-in-class ML scoring capacity for structure-based restricting affinity prognosis of protein-ligand edifices. The RF score determines the binding affinity and gives the entire model a performance score. The theoretical analogy represents how it is feasible to follow protein-ligand binding affinity from atomic perspective but also is executable for datasets with a definite view of accuracy expectancy. The training and validation test scores are mentioned as follows:

```
RF score train:  0.2638888888888889
RF score val:  0.125
```

4.4.10 Heat map analysis

Heat map is an information representation or visualization method that shows the magnitude of consequences as colors in two measurements and dimensions. The variety in shading might be by tone or hue, giving clear viewable signs to the analyzer about how the marvel is grouped or changes over space. There are two generally various classes of warmth maps: the cluster heat map and the spatial heat map. In a group heat map, extents are spread out into a lattice of fixed cell size whose lines and sections are discrete marvels and classes, and the arranging of lines and segments is purposeful and fairly discretionary, with the objective of proposing bunches or depicting them as found through factual investigation. The size of the cell is subjective yet enormous enough to be obviously noticeable. Conversely, the situation of an extent in a spatial heat map is constrained by the area of the magnitude in that space, and there is no idea of cells; the marvel is considered to differ constantly. Following is the heat map generated for the study.

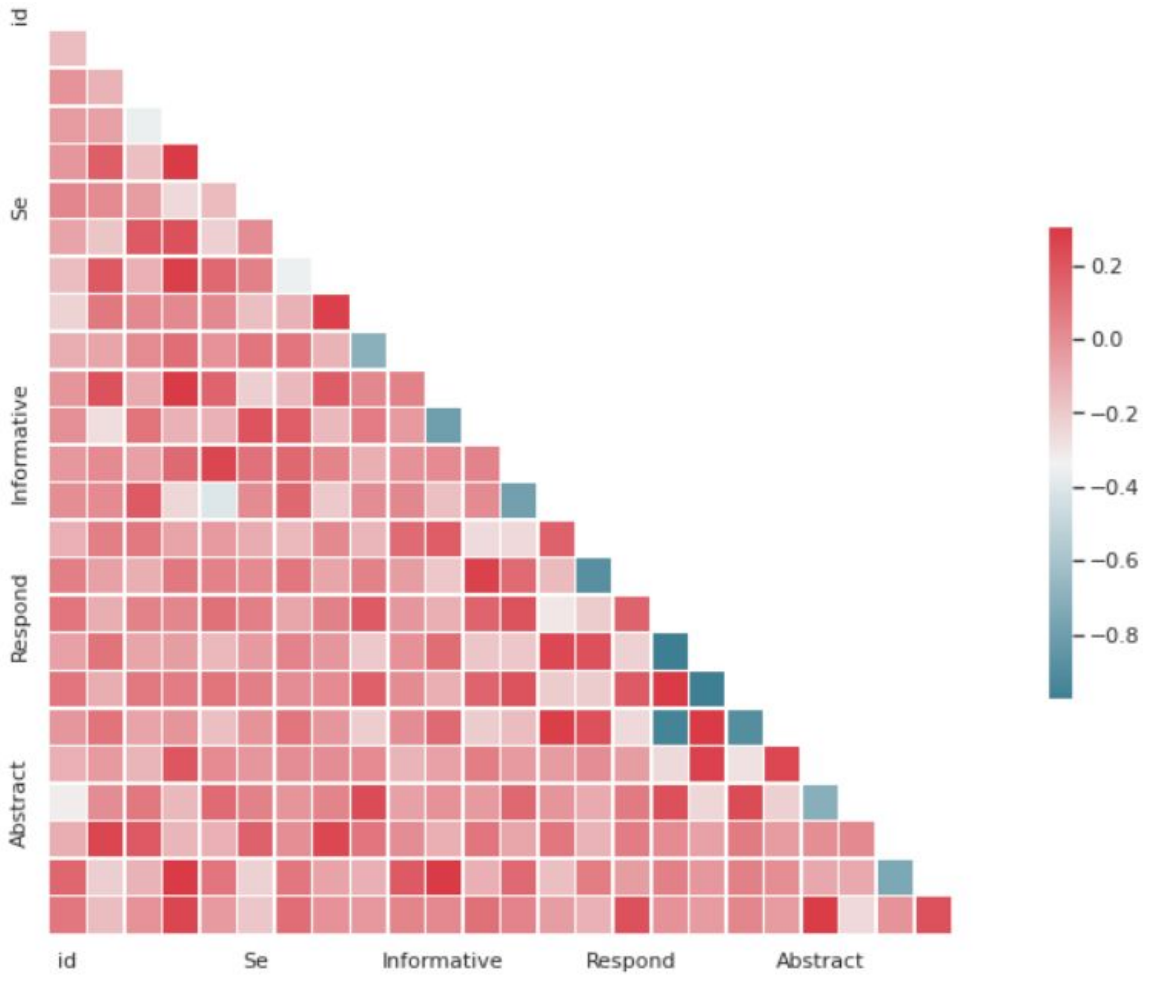


Fig. 4.11 Heat mapping of 5 traits.

4.4.11 Graphs and plots

The relational graphs that show the interdependence of features were optimized through seaborn and other visualization libraries to represent the feature importance with Archetypes and other traits in the study.

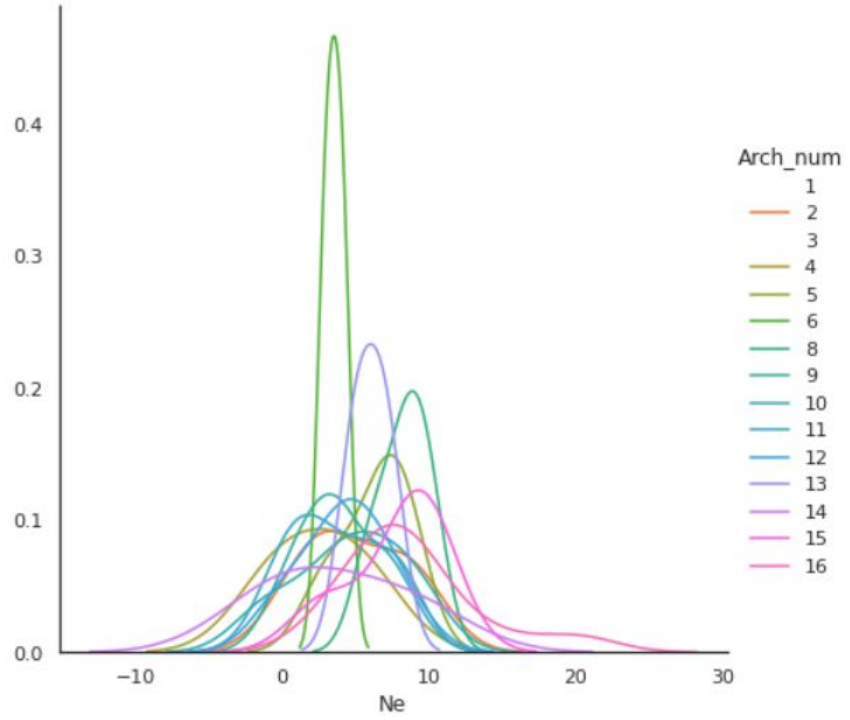


Fig. 4.12 Weight of ‘Ne’ for each Archetype.

The above-mentioned figure describes the significance of ‘Ne’ for all 15 types of Archetypes from the study.

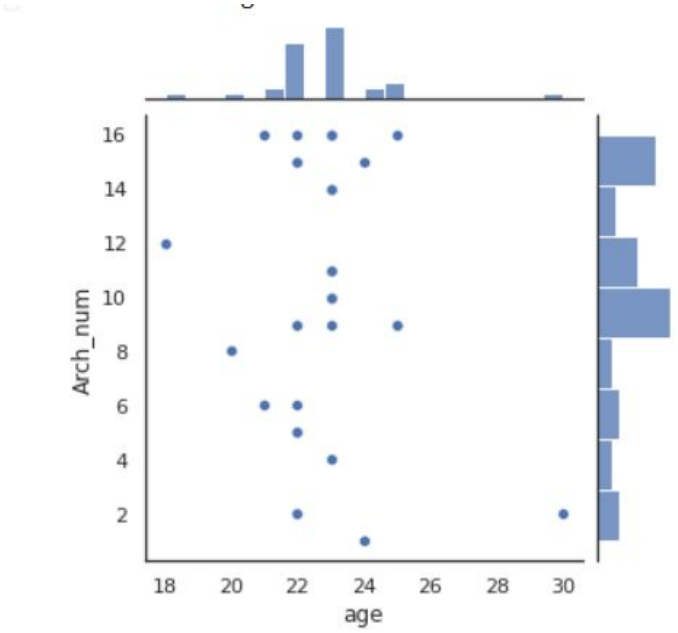


Fig. 4.13 Joint data plots Arch_num Vs age.

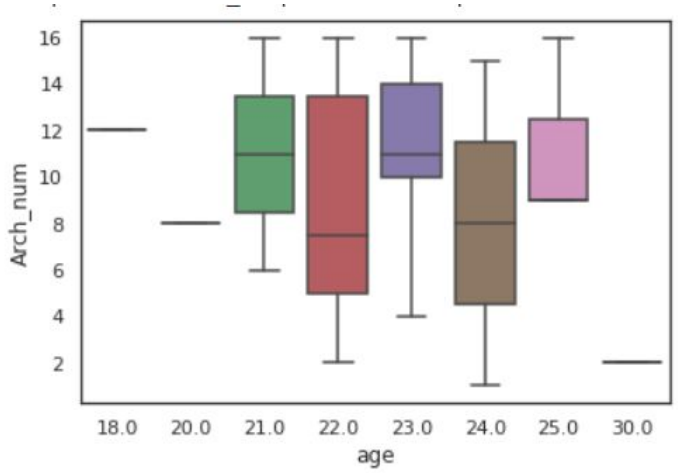


Fig. 4.14 Box plot for Arch_num Vs age

Chapter 5

Conclusion and future work

5.1 Summary

Our thesis project was about understanding human learning and decision-making processes based on Jungian analytical psychology using python and data science. The main purpose of our work was to ask the individuals we interviewed precise questions using Jungian analysis so that a better and more accurate result can be obtained than the tests that are available nowadays. The summary of our work is as follows:

- Primarily the focus was on designing and comprehending a questionnaire customized on the basis of Jungian Analysis. 84 people were questioned and 56 questions were asked to each of them. The concern was to create a dataset that will be integrated further into structuring and functioning.
- It's been observed that the archetype of an individual can be deduced better from the cognitive functions they use rather than the letters mentioned in the type indicator. That's the basis of Jungian analysis. Carl Jung first observed a pattern within human behavior that got more accurate and updated throughout the years.

- These cognitive functions are prevalent in every aspect of our lives. Starting from our behavior and reactions to our surroundings, to what we prioritize, how we affect one another as individuals and collectively, our interests and what drives us, what annoys us, and also how we interact with one another. These functions are a part of human nature, not nurture.
- K means clustering algorithm was used
- This thesis project is about conducting the test by asking certain questions and figuring out each participant's types based on the answer as these answers focus more on their cognitive features.
  This method gives a better understanding of human nature. Later k-means clustering algorithm and random forest classifier were applied to train our data for the system to understand the pattern.
- Libraries like pandas and sklearn were imported to formulate and analyze our data set respectively. Matplotlib and NumPy were used to generate grids and plot predictions.
- The result generation procedures consisted of labeling the centroids and the clusters. 2D and 3D cluster plots were generated and the graphical relation of WCSS vs Elbow Method was also obtained.
- The Random Forest Classifier was used to generate predictions of personalities with a supervised learning module. SImilar libraries were used for this analysis as well like pandas, sklearn. The “seaborn” library was introduced to have data visualization. A train set and a test set of equal entries were made. Data cleaning was executed as well as knn imputation. A Pearson correlation map was generated and split functions into training and validation.
- The results were generated in terms of feature importance, rf validation, and grid plots for the random forest model.

5.2 Limitations

Some limitations that we faced during our thesis work are as follows:

- Before the COVID-19 pandemic, the questionnaires were taken in persons and the answers were not limited to two or a few options.
  From those answers, it was easier to gather more accurate data. Because the functions can be deduced from such answers more easily as it's not just the answer they give but also how they give the answers and what they prioritize.
- After the pandemic, the interviews could not be taken face to face, and also it was not possible to get many people to take the interview because it is easier for people to avoid online links. Due to this less number of data, results may vary from ideal.
- It is observed people tend to answer more accurately if there are scenarios for them to relate to. As we had to take the interview online, we could not give them scenarios that were possible when we were taking physical interviews. Due to this reason, it may be possible that the interview takers couldn't answer some of the questions properly and gave answers based on their assumptions.
- People answered questions based on socially acceptable ways. Also, people like to answer questions in the way they would like to see themselves. This creates a bias called social bias.
- It was not possible to work with the column name, sex, profession, quadra, and archetypes for their categorical nature, hence it was chosen to introduce ids and archetype number which minimizes our possibility to generate results as characterized words.

- It was also observed that people like to answer questions in the way they would like to see themselves behaving. For example, an extroverted person might think he is introverted because he sometimes likes to be alone. But that's not the case always and that's why there remains a bias as well.

5.3 Future Possibilities

The findings from this study have important implications for further research and practice. Our project has a lot of potential if it is applied to the right field. Some of the many applications of our project are discussed below:

- Analyzing brain signal: This method can be used to analyze brain signal emission as similar types of people emit similar types of brain signals.
- Section division of preschoolers: Pre-schoolers can be divided into sections based on their personality types. This division will be efficient because similar types learn similarly. This section division will be helpful for the children than the traditional method of section division because teachers then can teach the children in different ways based on their learning types and this process can help to build up their base efficiently.
- Selecting suitable candidates for particular jobs: Our project can also be applied to select candidates for particular jobs as this method also reveals the decision-making styles of individuals. This will save some steps in assessment centers because some information about their decision-making styles will be available already.
- Jungian psychotherapy: this is a part of analytical psychology invented by Carl Jung. Jung stated that the collective unconscious was shared by all classes of people. Our study can help people in this therapy as this psychotherapy is intertwined with the interview taken by us.

5.4 Conclusion

This research has developed an efficient way to determine personality types than the frequently used methods. This project can help people in a lot of ways if it gets the opportunity to be implemented properly. Very few works have been done as our study incorporates psychology with data science. So as our project has some limitations, future research should be done on this and there remains a huge possibility that this work serves the general people in a lot of ways. Future research on these topics will be highly appreciated.

References

1. Understanding the Archetypes involving the eight functions of type (Beebe model), http://www.erictb.info/archetypes.html
2. Essential Terms & Theory, https://typeindepth.com/terms-theory/
3. What Are The Four Interactions Styles? | Structure: ESTJ, ESTP, ENTJ, ENFJ, https://csjoseph.life/what-are-the-four-interactions-styles-structure-estj-estp-entj-enfj/
4. What Are The Four Interaction Styles? | Starters: ESFJ, ESFP, ENTP, ENFP, https://csjoseph.life/what-are-the-four-interaction-styles-starters-esfj-esfp-entp-enfp/
5. What Are The Four Communication Styles? Finisher Types: ISTJ, ISTP, INTJ, INFJ, https://csjoseph.life/what-are-the-four-communication-styles-finisher-types-istj-istp-intj-infj/
6. What are the Four Communication Styles? Background Types: ISFJ, ISFP, INTP, INFP, https://csjoseph.life/what-are-the-four-communication-styles-background-types-isfj-isfp-intp-infp/
7. The Sixteen Personality Types, https://lindaberens.com/resources/methodology-articles/the-sixteen-personality-types/
8. Beebe, John – Jungianthology: A Podcast & Blog, https://jungchicago.org/blog/category/speakers-authors/beebe-john/
9. CS Joseph, https://www.youtube.com/channel/UCELhS3lbQQ8GVa2UeoVXAkQ
10. What Are The Four Sides Of The Mind?, https://csjoseph.life/what-are-the-four-sides-of-the-mind/
11. Linda Berens, https://www.youtube.com/channel/UCXBHBwndULvGq_Y7NNaRW0w
12. Welcome to CS Joseph Dot Life | CS Joseph, https://csjoseph.life/
13. cognitive-function-chart-2.jpg, https://www.nerdycreator.com/wp-content/uploads/2018/04/0513-cognitive-function-chart-2.jpg
14. cognitive-function-chart-1.jpg, https://www.nerdycreator.com/wp-content/uploads/2018/04/0513-cognitive-function-chart-1.jpg
15. personality-types-infographic.jpg, https://blog.adioma.com/wp-content/uploads/2019/05/16-personality-types-infographic.jpg
16. personality-types.png, https://www.prevuehr.com/drive/uploads/2017/08/16-personality-types.png
17. Analytical psychology, https://en.wikipedia.org/wiki/Analytical_psychology
18. Myers-Briggs 101: What are the Cognitive Functions? || MBTI 101, https://www.youtube.com/watch?v=fUOkMm43hGM
19. Constructing Survey Questionnaires; https://opentextbc.ca/researchmethods/chapter/constructing-survey-questionnaires/
20. Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco, CA: Jossey-Bass.

21. Logic Programming, Peter Norvig, in Paradigms of Artificial Intelligence Programming, 1992
22. A Survey on Machine-Learning Techniques in Cognitive Radios Mario Bkassiny, Student Member, IEEE, Yang Li, Student Member, IEEE, and Sudharman K. Jayaweera, Senior Member, IEEE, 2012
23. 10 Machine Learning Methods that Every Data Scientist Should Know, Jorge Castañón, 2019
24. 5 ESSENTIAL MACHINE LEARNING TECHNIQUES FOR BUSINESS APPLICATIONS; Oleksii Tsymbal, 2018; https://mobidev.biz/blog/5-essential-machine-learning-techniques
25. Craft, A. (2001). Neuro-linguistic programming and learning theory. *The Curriculum Journal, 12*(1), 125-136. doi: 10.1080/09585170010017781
26. Understanding Random Forest: How the Algorithm Works and Why It Is So Effective; Tony Yiu, 2019; https://towardsdatascience.com/understanding-random-forest-58381e0602d2
27. Celebi, M. E.; Kingravi, H. A.; Vela, P. A. (2013). "A comparative study of efficient initialization methods for the k-means clustering algorithm".
28. Cognitive class, k means clustering. Saeed Aghabozorgi, 2018