Evaluation Methods

Than Lwin Aung

Philosophy

Philosophy is a field of study to explore the existential and essential criteria of humanity and nature.

There are three primary branches in Philosophy, namely, **Ontology**, the study of Existential Reality, **Epistemology**, the study of Knowledge, and **Axiology**, the study of Value.

However, as my research area is primarily concerned with Evaluation Methods, I will mainly focus on Axiology, the study of values.

Axiology

The highest level of knowledge, as far as I concern, is the **knowledge of values**, without which making **good decisions** is almost impossible.

There are 4 primary premises, which can intrinsically determine the values of some existential reality: the **truth**, the **good**, the **just** and the **beauty**.

In essence, making a decision is just selecting options, which contain one or more of those 4 primary premises.

Actually, the study of values has become more important as the agency of machines has become more prevalent than ever before.

So, what really is the “**value**” of an object or an experience?

Values

When it comes to “Values”, there are many definitions of values; however, I will only look into “**Intrinsic Values**”.

Intrinsic values are the values which can be determined without any reference to externalities.

However, **Extrinsic values** or **Instrumental Values** are the values which are dependent on externalities, such as environments.

Therefore, when I say “value” of something, it is only with respect to its Intrinsic Values.

Evaluations

To understand knowledge of values, some forms of **Evaluation Functions** are necessary.

The easiest understandable Evaluation Function is “**Ratings**”, which is part of the well-established norms of evaluation, called **KPI**, Key Performance Indicators.

For example, what is the rating of “Titanic Movie”? If we give the rating score 9.5 out of 10, what does this score really mean?

More importantly, how do we reach to this conclusion that “Titanic Movie” worth the score 9.5 out of 10? Which factors actually determine this rating?

Mathematical Evaluations

Now, let us look into one of the Evaluation Functions, for which we could employ mathematical evaluations to determine the values of a set of objects of existence.

At this point, it should naturally occurs to us as to how we could mathematically determine the intrinsic values of an object or an experience.

Before looking deeply into Mathematical Method of Evaluations, we will look into "**Sentiment Analysis**", along with **Similarity Measure**.

Sentiment Analysis

Since the early days of Natural Language Processing (NLP), Sentiment Analysis has been widely used to determine the **opinions** of the public with respect to certain **topics**.

Even the widely available social media employs "**Reaction Buttons**" to survey the opinions of the public.

Document and Sentence Level Sentiment

	Score	Magnitude
Entire Document	0.46	1.531
In face of hardships, he is relentless.	0.805	0.915
His couregous demanor has been tempered by the years of overcoming the failure after another.	0.115	0.616

Score Range	0.25 — 1.0	-0.25 — 0.25	-1.0 — -0.25

Sentiment Analysis

In a sense, we can say that Sentiment Analysis is one of the mathematical evaluation methods, which give us a **Scalar Score** of the **Distributions of Topics** in a paragraph or a sentence.

Therefore, we can conclude that the **Sentiment Score** of a Paragraph is the **Intrinsic Value** of the Paragraph determined by the Mathematical Evaluation Function.

Actually, in addition to text, we can use **Sentiment Score** for any **Modality** of Data, without any reference to Externalities.

Thus, **Sentiment Score** is the Mathematical Measure of **Intrinsic Value** of the Modality of Data. To put it more simply, it is the measurement of “how much score (value) do you want to give to an object or an experience”?

Scores

In a broader sense, “**Scores**” are Mathematical Quantization of Intrinsic Value of Objects, which is the Cornerstone of KPI.

Even in our everyday life, we will inevitably encounter one score after another, from Exams through Social Media to Customer Services.

Now, let us look into Mathematical Evaluation Methods, from the Perspective of **Tensors**, and also why **Linear Forms** can be considered as **Evaluation Functions**.

Linear Forms

Suppose we have a vector space $\mathbf{v} \in \mathbf{V}$ with the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n\}$. Then, we can represent a vector **v** as a Linear Combinations of the Basis $\mathbf{e_i}$ such that:

$$\boldsymbol{v} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_1\boldsymbol{e}_3 + \cdots + a_n\boldsymbol{e}_n$$

We can actually define a Linear Function $l\colon \boldsymbol{V} \to R$ such that:

$$l(\boldsymbol{V}) = V^*$$

$$l(e_i) = a_i$$

And such Linear Function $l\colon \boldsymbol{V} \to R$, is known as **Linear Form** or **Linear 1-Form**.

Actually, Linear Form is also a Linear Function between Vector Space **V** and Dual Space **V***.

$$g_i^*(\boldsymbol{v}) = v_i$$

Example of Linear Forms

Suppose we have a vector space **v** ∈ **V** with the basis {**x, y, z**}. Then, we can represent a vector **v** as a Linear Combinations of the Basis {**x, y, z**} such that:

$$\boldsymbol{v} = a_1\boldsymbol{x} + a_2\boldsymbol{y} + a_1\boldsymbol{z}$$

Then, Linear Function $l\colon \boldsymbol{V} \rightarrow R$ is defined such that:

$$f(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = R = a_1x + a_2y + a_1z$$

If (x, y, z) are the co-ordinates (**unit direction**) such that {(1, 0, 0), (0, 1, 0), (0, 0, 1)}, then

$$f(1,0,0) = a_1$$
$$f(0,1,0) = a_2$$
$$f(0,0,1) = a_3$$

Such Linear Function is also called **Evaluation Function**, because it can tell us something about vector **v**. If we provide a **directional coordinates**, it will give us the **components of vector in that direction**, which is extremely important to analyze vectors.

Evaluation Function

Therefore, the Linear Form l(v) can be considered as Evaluation Function.

Suppose, vector space $\mathbf{v} \in \mathbf{V}$ with the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n\}$, which represents the letters of a language, such as {“a”, “b”, … , “z”}.

Then, any word can be considered as a Linear Combinations of such basis letters, such that

$$\text{is} = v = 0a + \cdots + 1i + \cdots + 1s + \cdots + 0z$$

Where **1** and **0** are dual basis of Covector, while {“a”, “b”, … , “d”, … , “z”} are basis of Vector.

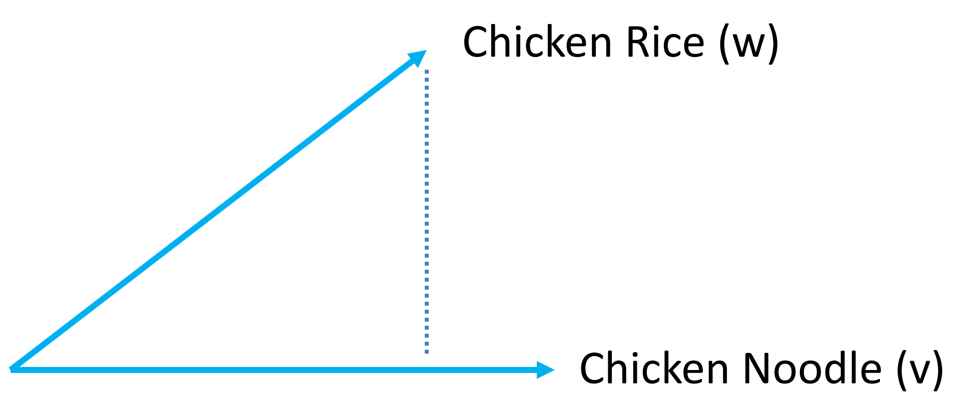
Then, what Linear Form l(v) can represent is the score of the word or the intrinsic value of the word.

Inner Product Space

If the Linear Form l(v) can be considered as Evaluation Function, then Inner Product l(v, w), can represent the evaluation score of the two words, and more specifically, it measures the similarity score between them.

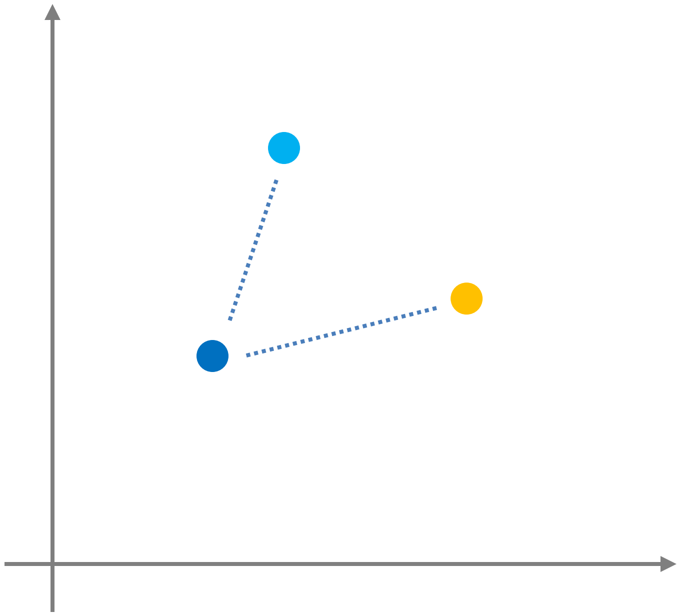
In other words, we can consider l(v, w) as vector **v** is evaluating vector **w**, with respect to vector **v**.

For example, what is the similarity between “Chicken Rice” and “Chicken Noodle”, based on the intrinsic values of word?



K-NN

K-NN (K- Nearest Neighbor) is the one of the evaluation methods to find the distance of similarity in Inner Product Space between vector **v** and **u**, which is based on the notion that similar objects have similar intrinsic value.



Dual Space of Vector and Covector

As we can see, **Vector Space V** $\in \mathrm{R}^n$ can represent the **Features of the Objects** while **Covector Space V*** $\in \mathrm{R}$ can represent the **Value of the Objects**.

For example, if we consider a movie (motion picture) as Vector **V** $\in \mathrm{R}^3$ (x, y, t), then **Covector Space V*** $\in \mathrm{R}$ of that movie can represent the Rating of that overall Movie.

Therefore, if we compare two movies, we are not comparing the two movies frame by frame; instead, we are comparing their intrinsic values in **Covector Space**.

Pleasantness: 0.21

Pleasantness: 0.65

Tensor Fields

Now, we can extend the **Vector Space** and **Covector Space** to **Vector Fields** and **Covector Fields**, and ultimately, they are called "Tensor Fields".

When it comes to Fields, we use **Differential Form**, instead of **Linear Form**, and we use **Interior Product** Space, instead of **Inner Product** Space.

In a sense, **Covariant Derivative** can be considered as the measurement of similarity between vector Field **W** and **V** in their Interior Product Space.

$$\nabla_v W = \nabla_{v^i \boldsymbol{\partial}_i}(w^j \boldsymbol{\partial}_j)$$

In other words, we are measuring the similarity of Vector Field **W** to **V**, along the flow of Vector Field **V**.

Measurement of Scores

Why the measurement of scores by mathematical evaluation functions are important is we could possibly measure the Intrinsic value of an object in this way.

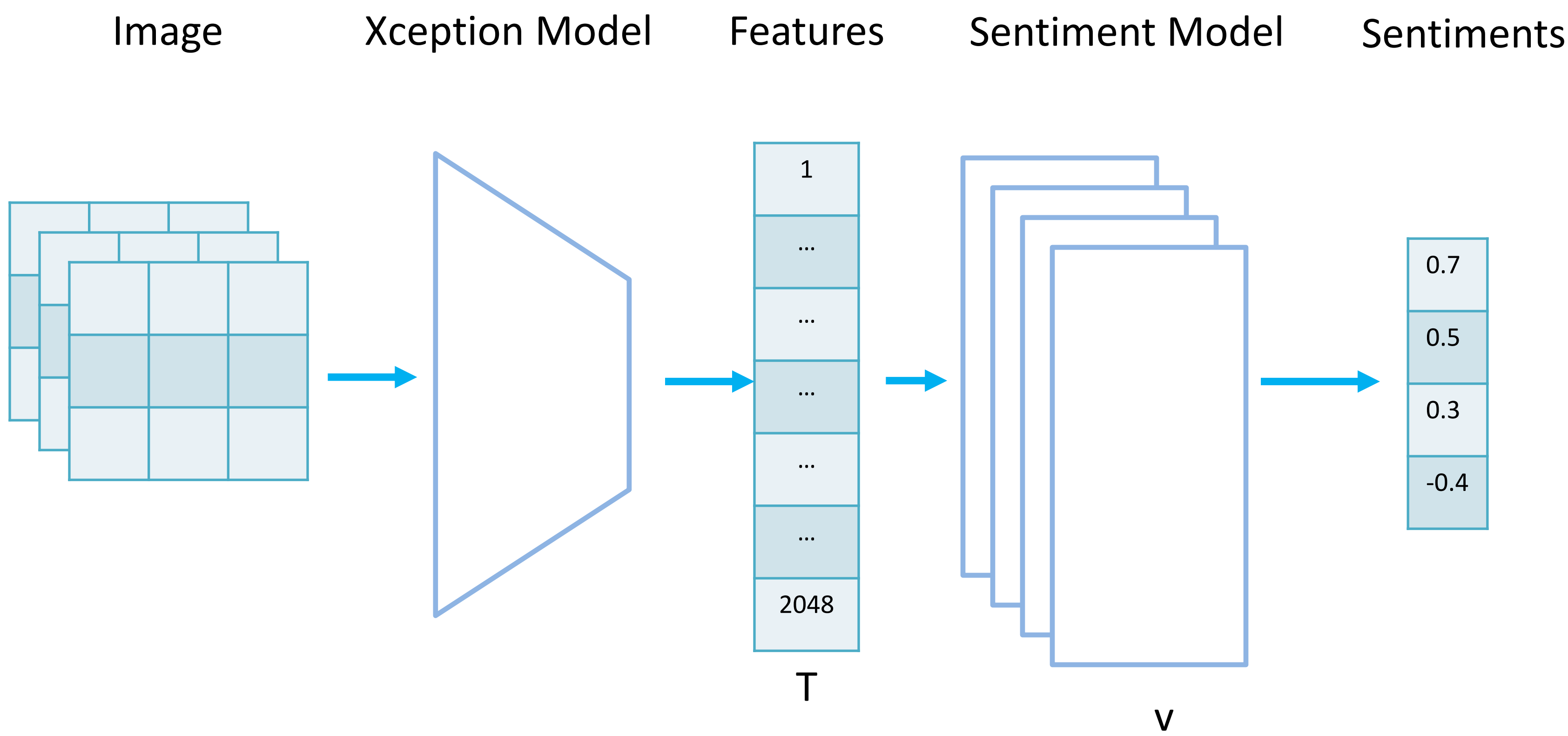
However, we need to be aware that depending on evaluation functions (**V**), we can get different scores even for the same Intrinsic Value (**W**).

In addition, with the scores of the Intrinsic value of objects, we can use them in Information Retrieval in Memory, along with Decision Analysis.

However, it still begs for more questions as to what define the generic sentiment values of objects, such as human emotions, in terms of Tensor Fields.

To answer this, I would propose a **novel solution** in which Sentiment Model is considered as an Evaluation Function.

$$L_v T = L(v, T)$$



Sentiment Evaluation is based on the concept that a Tensor **T** is evaluated in terms of a vector **v** in **Sentiment Space** (V3A), which would give us a **sentiment score**.

```
# Input with shape of height=299 and width=299
inputs = Input(shape=(2048), name="image")
labels = Input(shape=(4), name="label")

dense1 = Dense(512, activation='tanh')(inputs)
x1 = Dense(1)(dense1)

dense2 = Dense(512, activation='tanh')(inputs)
x2 = Dense(1)(dense2)

dense3 = Dense(512, activation='tanh')(inputs)
x3 = Dense(1)(dense3)

dense4 = Dense(512, activation='tanh')(inputs)
x4 = Dense(1)(dense4)

y = layers.Concatenate()([x1, x2, x3, x4])

output = SIMPLE_LOSS(name="loss")(labels, y)

model = Model(inputs=[inputs, labels], outputs=output)
```

Sentiment Scores

As we can see, each Dense Layer (MLP) can be considered as **sub manifold** of **the coordinate function** φ such that:

$$\varphi: M \to N \to R \text{ , where M = 2048, N = 512, R = 1}$$

$$x_j = \varphi_j(T) \text{, where T = Input Tensor}$$

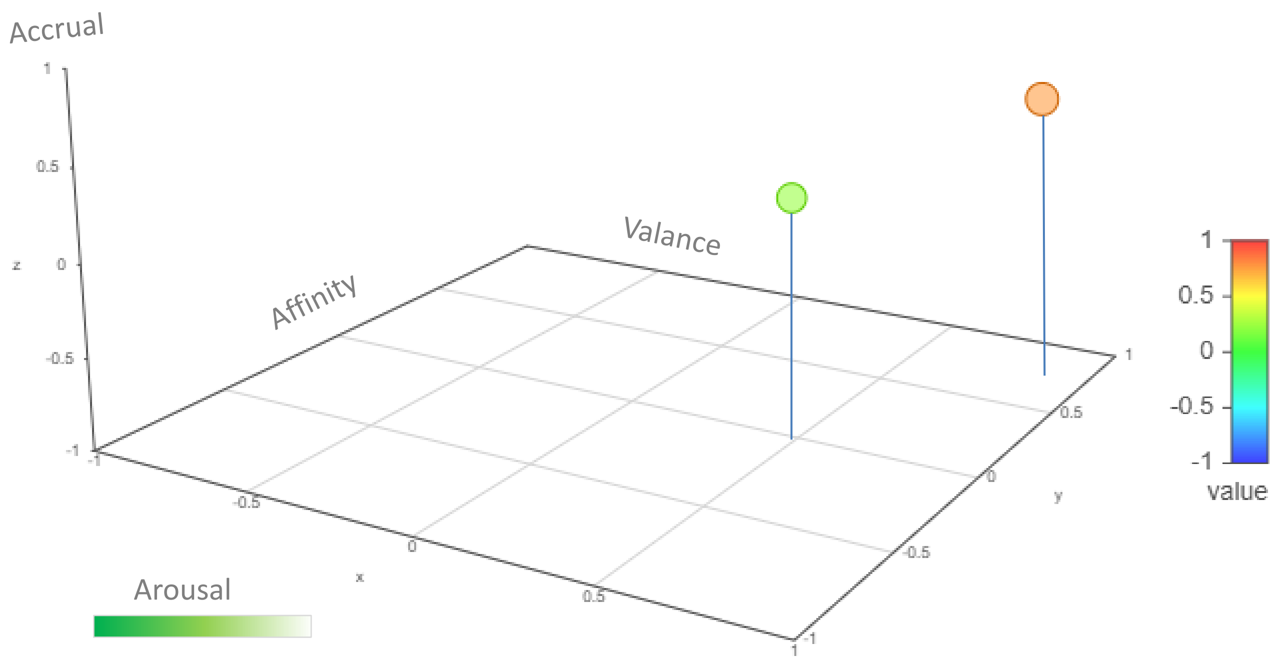
$$\boldsymbol{v} = \sum_{j=1}^{k} x_j v_j = [x_1 v_1, \cdots, x_k v_k] \text{ , where } \mathbf{v} \text{ = Output Tensor}$$

Therefore, $$\boldsymbol{v} = \sum_{j=1}^{k} \varphi_j(T) v_j$$

As we can see, **v** is a vector in Vector Space with k-dimension. However, I will define such vector space as an Equivalent of Sentiment Space.

Sentiment Space

The sentiment space is defined as k = 4 Dimensional Vector Space, which is a Linear Space of 4 bases : { Valance, Arousal, Affinity and Accrual }.



Sentimental Evaluation

As we can see, we can define Sentiments as Sentimental Evaluation in Vector Space. To put it more simply, when a Tensor **T** is projective onto a Vector Space **v**, its covectors can represent evaluation scores.

From this, I have reached to the conclusion that covectors can represent "values" while vectors can represent "features". In other words, "features" and "values" could be the dual space of the same reality, defining **semantical** and **sentimental** aspects of reality, respectively.

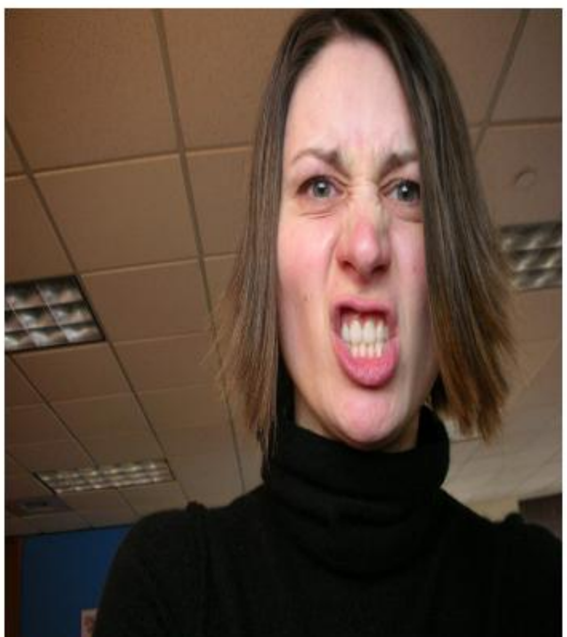
Everyday, our eyes can see the different objects, along with different visual experiences. However, we need to evaluate each visual experience with respect to our survival and social scores.

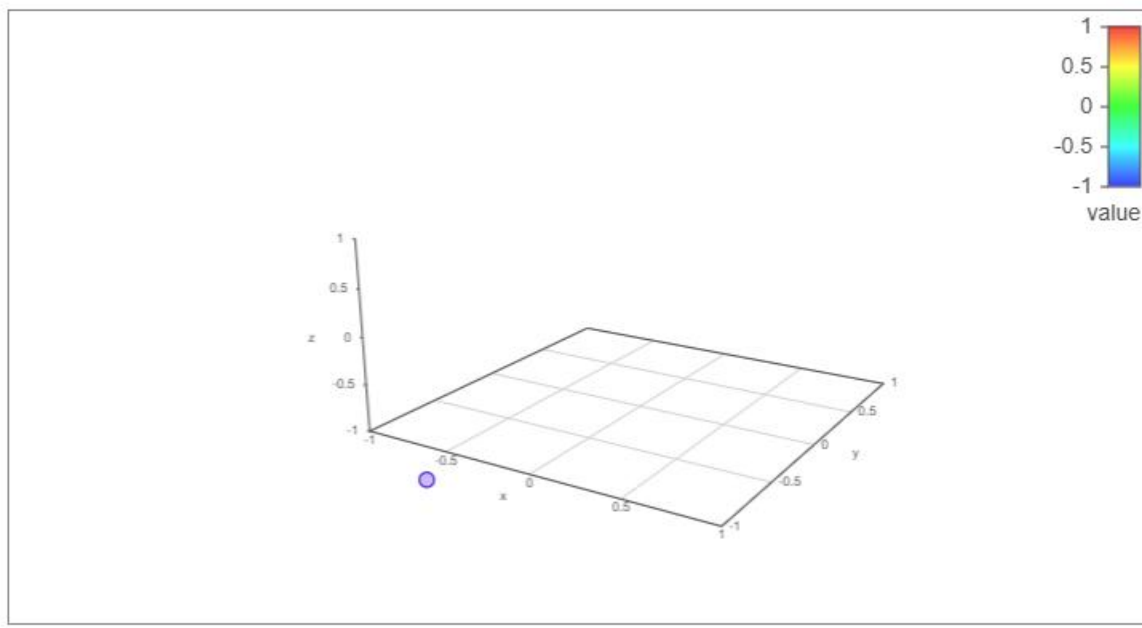
How do we feel when we see something? Some will be evaluated as "good" while others will be discarded as "bad", as the ancient wisdom said, "when we see something as beautiful, the others become ugly".

Experimental Outcomes

We will look into the experimental outcomes to support my hypothesis of Sentimental Evaluation.

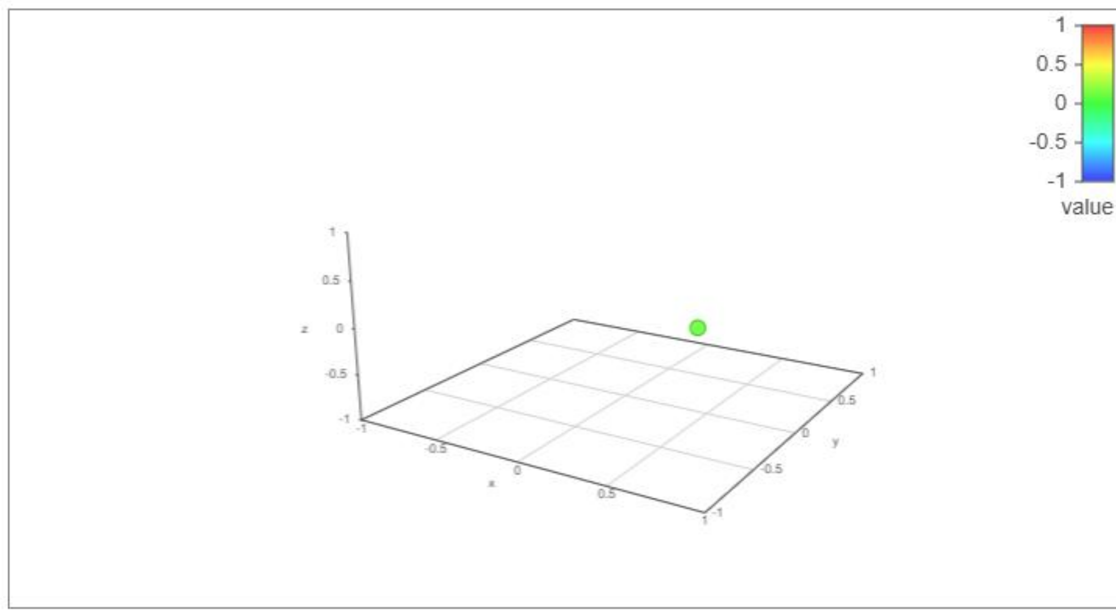
I am feeling:Hatred,Fear,Shameful,Depressed,Guilty





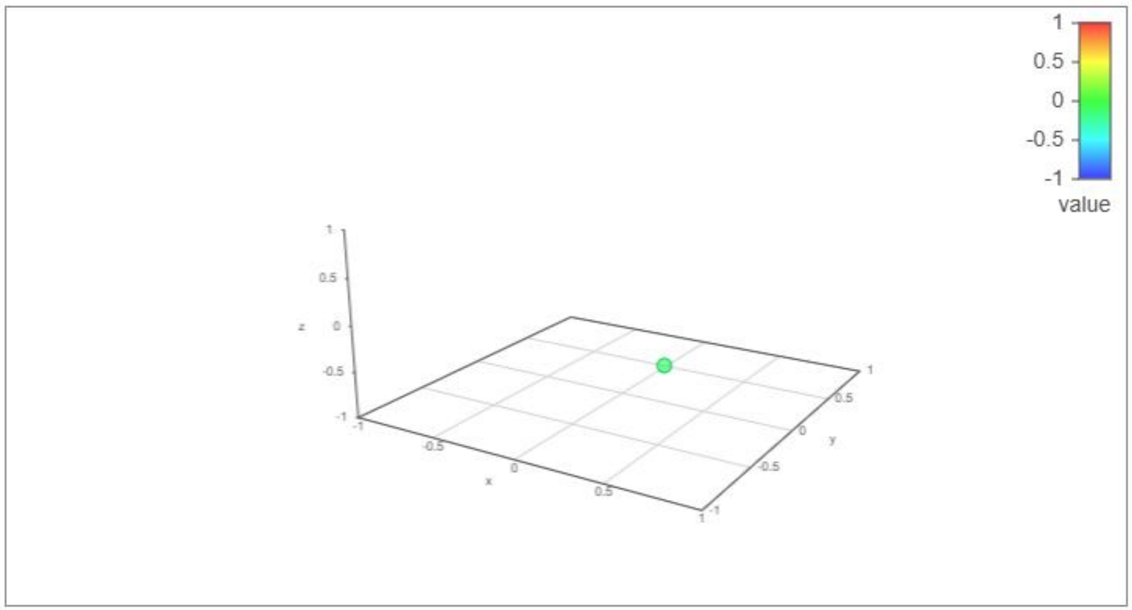
Experimental Outcomes

I am feeling:Curious,Cheerful,Enthralled,Happy,Like



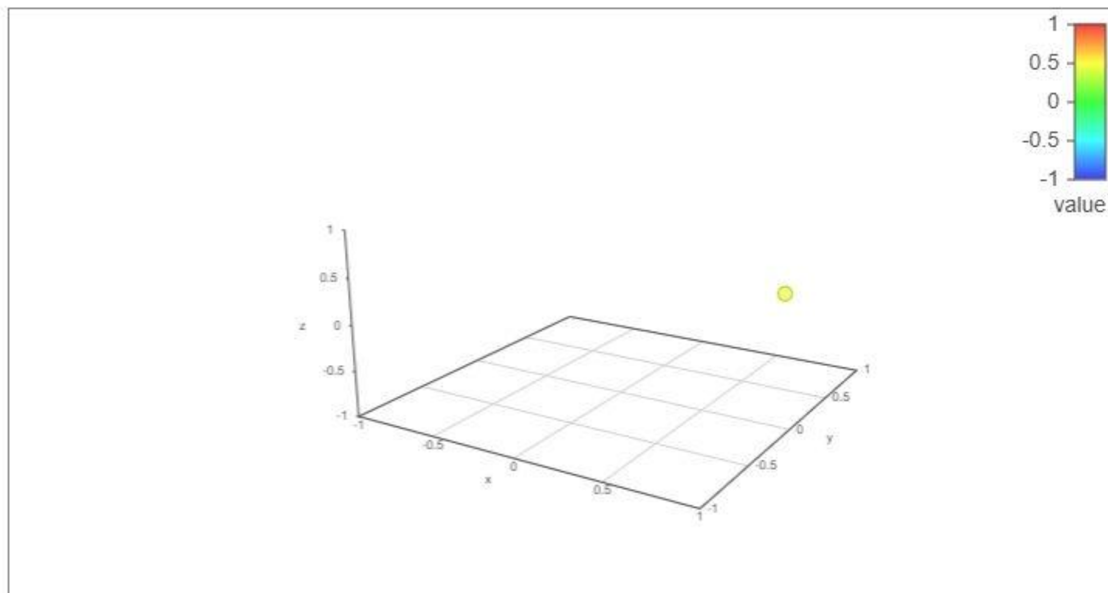
Experimental Outcomes

I am feeling:Missing,Jealous,Nostaligic,Indulgent,Anxious



Experimental Outcomes

I am feeling:Tempted,Compassionate,Love,Amused,Enthralled



Similarity in Sentiment Space

As we can see, we can define **Similarity in Features**, and we can also define **Similarity in Sentiments**.

Although different images are not **semantically similar**, they are **sentimentally similar**.

Therefore, with Similarity in **Semantical Features** and **Sentimental Values**, we can implement better **Retrieval Systems**.

Similarity in Sentiment Space