Affective Computing

Dr. Abhinav Dhall
Department of Computer Science and Engineering
Indraprastha Institute of Information Technology, Delhi

Week - 06
Lecture - 01
Emotion Analysis with Text

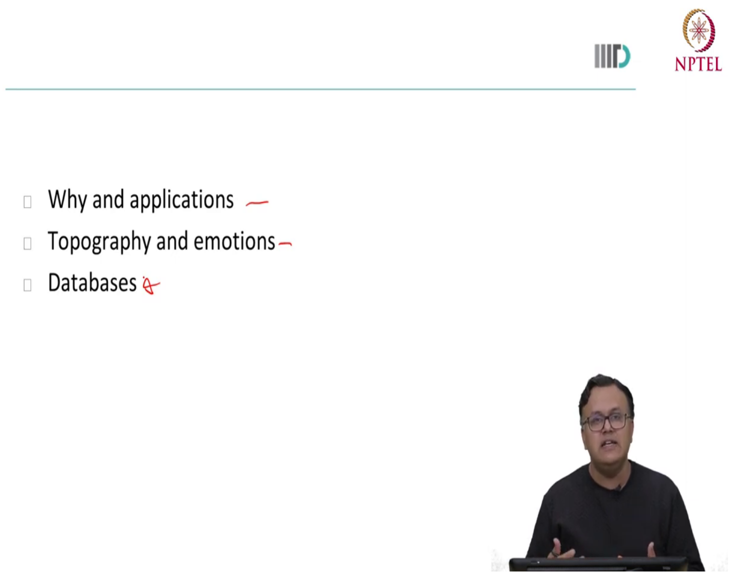
Hello and welcome. I am Abhinav Dhall from the Indian Institute of Technology Ropar. Friends, today we are going to discuss about analysis of emotion of a user through the text modality and this is part of the ongoing lecture series in Affective Computing. So, till now we have discussed how we can use the voice signal of a user, the facial expressions, the head pose of a user to understand the emotion which either the user is feeling or the one which is perceived.

Now, there is another modality which is very commonly used in the affective computing community that is the text and the reason is quite obvious. We see so many documents around us. We see conversations happening on chatting platforms. So, you would have seen billions of users on platforms such as WhatsApp, Telegram and so forth and they are conversing with text.

So, how can we understand the affective state of a user when the communication medium is text? And on contrary to the analysis with faces and voice what is really interesting to note with something with text is that text is not just about what is being communicated by a user in real time, but it is also about let us say looking at the emotion conveyed in chapter of a book, looking at the emotion conveyed by a particular comment which was posted few months ago on a website.

So, both online and offline text is being analyzed for understanding the emotion in different use cases.

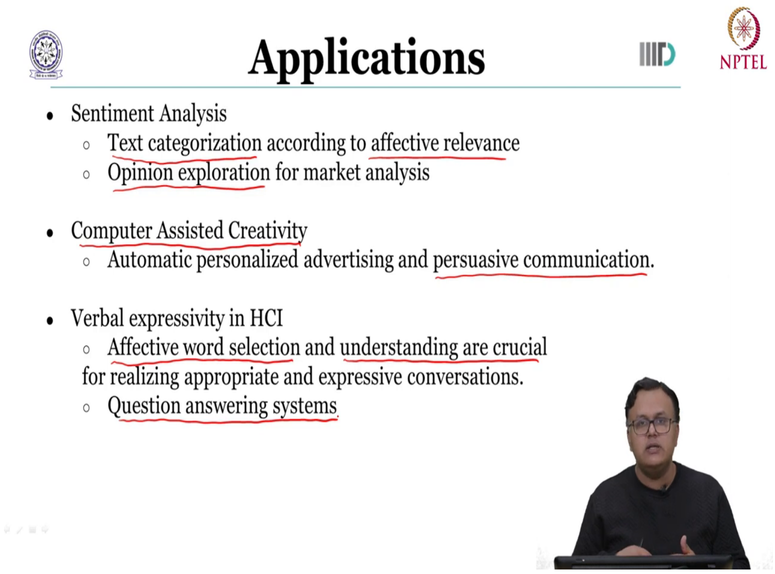
(Refer Slide Time: 02:24)



So, moving forward in this lecture, first we are going to talk about why text is important as a modality for understanding the affect of a user. Then we are going to look at some of the applications. From this we will go a bit back in time. So, we will explore how emotion has been conveyed through topography, how different fonts are used to convey the different emotions to a user when he or she is reading certain text.

And then we will move on to the databases. So, we will look at some of the resources which are available in the affective computing community where text has been labeled for its perceived effect.

(Refer Slide Time: 03:15)



Now, when we are typically looking at a conversation. Now this conversation let us say is between a user and a virtual avatar. So, we can have the camera modality, we can have the microphone, right. Now, when you are recording what the person is saying, you can also use speech to text in this pursue and understand the lexical content.

What is being said that conveys the emotion, how is it being said, is the communication more metaphorical or there are some implicit meanings which are being conveyed, right. So, there are several applications which are based on this construct of how emotion is conveyed through text.

Now, typically when we talk about emotions in text, friends you will see that in literature sometimes the keyword emotion and sentiment could be used interchangeably. However, this is a very simple yet extremely important difference. Emotion is what a user is feeling. So, either the user can tell his or her emotional state or what is being perceived by another person who is communicating with the user. So, the perceived emotion.

Now, with respect to text, let us say you are reading a comment about certain product on a website. Now, in this case, we are interpreting the affect conveyed by the written text in the comment as a third person. Now, this is the sentiment. So, what is the sentiment which a particular lines or text that is communicating to the user who is reading that text?

Now, we see that this type of sentiment analysis in text that is based on the text categorization according to what is the affective relevance. So, you read a comment on a website regarding a product. After you finish reading, you can actually tell if the user who was writing his or her feedback about that product felt positive or negative about that product, right.

Further, this is not only limited to the comments on these platforms, online platforms, but also it is very commonly used in understanding of the opinion for markets, right. Let us say a news article comes in about a certain company, about a certain product. Now, what is the sentiment which is conveyed in that news article?

And with that, we can try to understand that let us say what is the generic public opinion about that particular context regarding which that news article is written. Further emotion in also is commonly used in computer assisted creativity. Now, an example of that is, personalized advertisement that can be created by understanding the affective state of the user and the same can be now conveyed in the form of text to the user.

So, that the appropriate emotion is conveyed in this text. Now, this can lead to a more persuasive communication with the user, right. Let us say the text which you read in an image based advertisement that is using the words which are conveying let us say the excitement which a product could bring to a user, right. So, that can be conveyed within that text of how the framing of the text is done. Further for adding expressivity in human computer interaction, we use the emotion which is reflected through text.

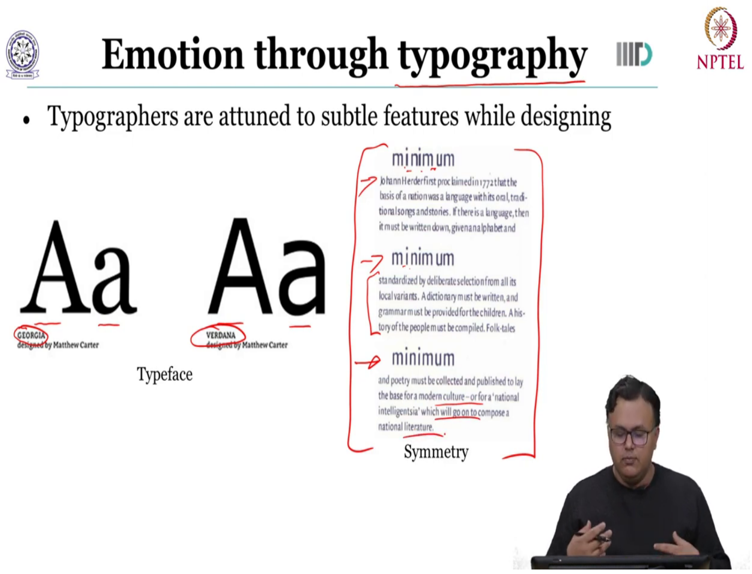
What that means? Let us say a user is interacting with the machine. The machine is going to give a verbal feedback. Now, this verbal feedback would of course, be a speech to text conversion. Now, this speech to text conversion would from the user would give the content about the lexical knowledge of what the user was saying. And now when the machine has to give this verbal feedback, it will do text to speech synthesis.

But what it means the appropriate text which needs to be replied back to the user that can have the emotional words added to it so as to convey the appropriate emotion, right. So that means, the word selection for conveying the correct affect. And then of course from the understanding perspective from speech to text when the user says is extremely important.

And in both these cases the text based affect analysis comes into the picture. You need to pick up the right words to convey to the user. You need to understand the text which the user spoke to you. The same is also friends applicable in a use case such as the question answering system. So, how to convey the emotion behind let us say answer which a machine is giving to a user, right.

So, that would be based on the word selection and so forth. That means your system needs to pick up the right words which will convey the right emotion.

(Refer Slide Time: 09:31)



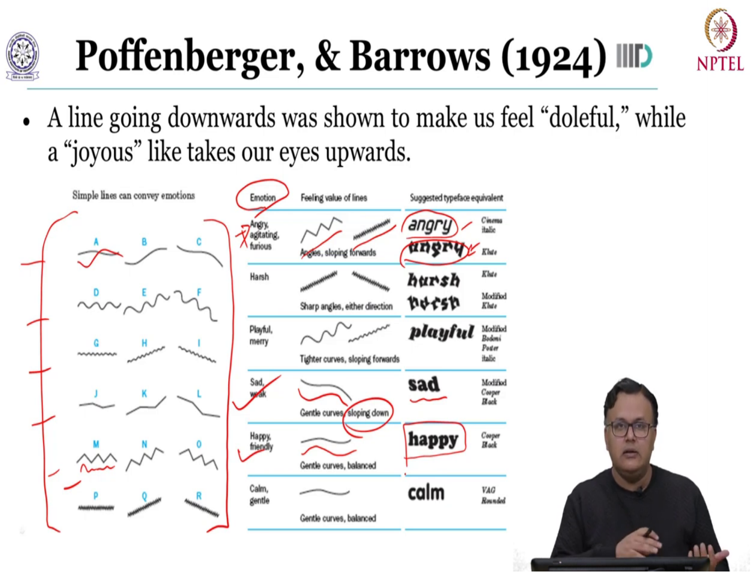
Now, from this let us move to how emotion has been conveyed in the past decades in the field of typography, ok. Now, typically typographers they would use these minute information to convey how the content is interpreted in terms of the emotion which is supposed to be conveyed by a text.

Now, friends if you notice the alphabets uppercase A and lowercase a in two different fonts Georgia and Verdana, right. You can actually see here that the way these characters are represented in these two fonts is actually going to in a very subtle way convey the affect which is supposed to be understood by the user. This not only is limited to the shape of the characters, but also to the placement as well.

Now, observe this example here. In this case you notice that there is difference in the gap between the alphabets. So, if one wanted to read this stanza it would actually of course, not only be a bit difficult to read it out, but it will be non trivial to understand sometimes the affect which is conveyed. Now, if we improve this in the second example wherein at least the gaps are more standardized we see that the reading and then further the interpretation that improves.

Now; however, when we are introducing symmetry with respect to these gaps in the third case notice how trivial it is to read and poetry must be collected and published to lay the base for a modern culture and so forth, right. So, the emotion that is now conveyed more clearly and the content is also understood by the user.

(Refer Slide Time: 11:55)



Now, there is this very interesting work which was proposed in 1924 by Poffenberger and Barrows. What they are saying is the way in which a line is drawn, the style of the line would represent the emotion which is supposed to be conveyed by let us say a word which is written by using the style of those lines.

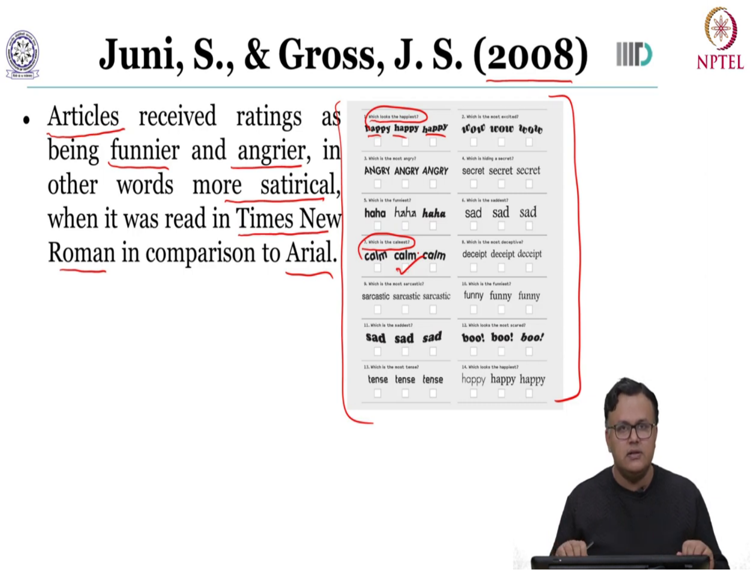
Now, notice this friends. So, here you have different styles of how lines can be written you know for example, A to R here are the different styles. Now, with the curvature and the frequency of change of the line you could have let us say smooth lines or you could have these zigzag seesaw pattern lines.

If you use that you can convey different type of emotions, ok. For example when we wanted to convey angry furious kind of emotion we can use these seesaw angular sloping forward lines and when we use these lines these shapes and then we create these words. For example, here you see angry you can actually understand that the intensity of the emotion which is supposed to be conveyed by that word it varies in these two different iterations.

So, of course, this is very subjective I would say that the second iteration of angry that to me is far more intense as compared to the first iteration where we are using more smoother lines, right. So, if you are using these angular patterns, you are able to convey far more intense angry react emotion.

Now, let us pick up the sad emotion, right. So, in this case if you have a gentler curve which is sloping down and then you use these suggested typeface equivalent, you can actually observe that this is slanting backwards sloping down. So, text in this would be perceived a bit more seriously in the gravity of the content which it is conveying.

(Refer Slide Time: 15:26)



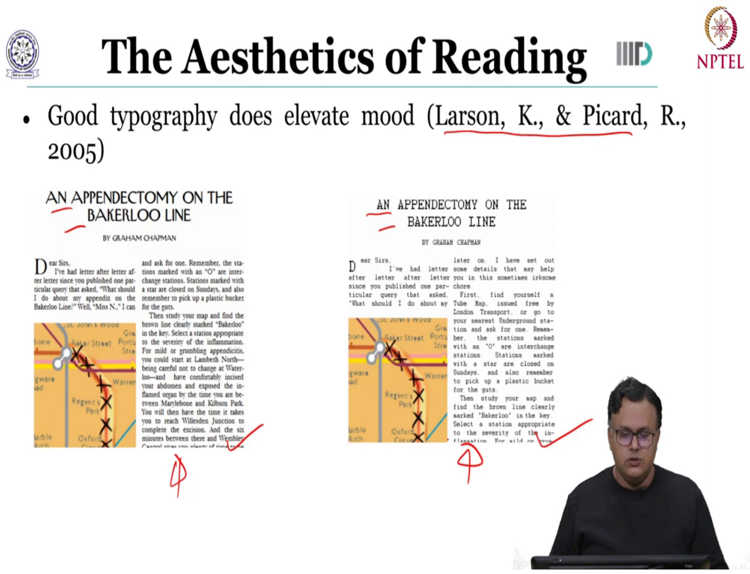
Now, if you were to compare that with happy friendly then you know it is observed that if you have these gentle curve balance lines which are not sloping down like sad then this is more easily conveyed, right. So, just by changing the typeface the orientation of the basic components which are coming in together for the word formation.

We can have the user interpret the emotion from the text easily and this is how you can convey the emotional content clearly to the user. Now, let us look at a bit relatively recent work. So, in this Juni and Gross in their work from 2008 they got a set of (Refer Time: 15:40) who added ratings, ok. Now, these ratings were given to articles as being let us say funnier and angrier or in other words more satirical. And they wanted to compare Times New Roman font with the Arial font.

Now, this kind of questionnaire was given and what you can actually see here is, let us say when we are talking about the question is which looks the happiest and you see these three options, right. So, of course, this is subjective, but the orientation in which the text is written which is of course, the font style that essentially was given to have effect on the user.

So, that means, the same word would be conveyed a bit differently in its emotional content. So, they found that you know in for example, in this case when you look at calm, the word calm itself is better conveyed in the second version because of the stability which has there in the alphabets, right.

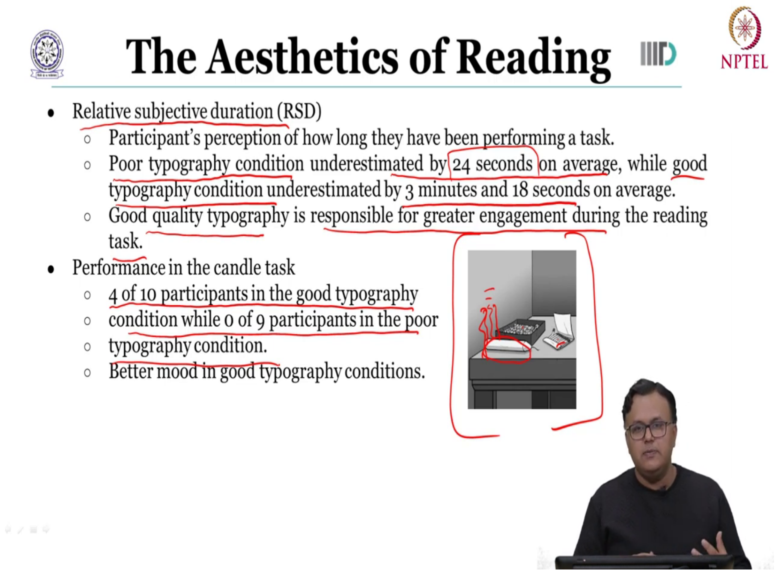
(Refer Slide Time: 16:58)



Now, let us look at another very relevant interesting work. Now, in this work Larson's and Picard they wanted to understand the aesthetics and its effect on reading, ok. Now, what do we have here? We have two versions of the same text. And if you close you observe they are in different font styles.

What they found was that you know if you have a good typography then it should effect the perceived affect and the real affect of the person, right. So, it could elevate the mood when you are reading let us say this article as compared to the same article in a different style.

(Refer Slide Time: 17:52)



So, to this end they looked at two task. The first is friends the relative subjective duration RSD. So, in this the participant's perception of how long they have been performing a task was evaluated. What was seen is when you have poor typographic conditions the participants, they underestimated the duration of the task by 24 seconds on the average.

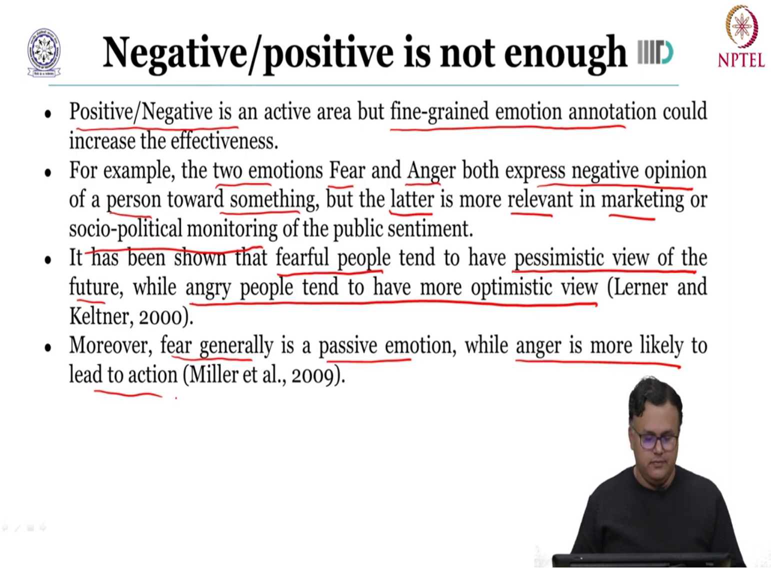
So, let us say they have been performing a task for n seconds when the text which they were reading as part of the task they were given with poor typographic conditions. They said well you know the duration for which I have been working on this task is roughly 24 seconds less than the actual duration. Why? Because In this they had more cognitive load because of the poor typography.

However, when good typography condition based text was presented they underestimated the duration by an average of 3 minutes and 18 seconds which means the reading was far more easier, right. And that is how the time the perception of time passing by that was faster. Now, what we learn is that good quality typography is responsible for greater engagement during the reading task, the user is more engaged and it is a more immersive experience.

Now, they also did a candle task. So, this is a very old cognitive task from the 1945 proposed by Duncker. So, what the task is, you have a candle, you have the match sticks, you have some pins. What you want to do is, you want to take this candle, you want to place it on the wall somehow that when you are going to burn that candle the wax should not fall on the table, right.

Now, in this particular case from the studies participants 4 of the 10 participants were in the good typography so they read the instructions. While 0 of 9 participants in the were in the poor typographic conditions. Now, of course, the participants who were given the same content in good typographic conditions they performed better in this case.

(Refer Slide Time: 20:32)



Now, with respect to the conditions of the how emotion is presented, I have already used in the beginning terms you know the positive emotion or negative emotion which is perceived let us say from a comment. But in a lot of cases that when you are looking at the text this might not be enough, ok. So, we see that in the works looking at text modality for emotion positive and negative is very commonly used.

But fine grained emotion annotation that even though it is more effective is lesser used in the works of course, there are the very obvious reasons right, data annotations and so forth. Now, if we take an example friends, for two emotions fear and anger. Now, both of this are expressed negative opinion of a person let us say towards something.

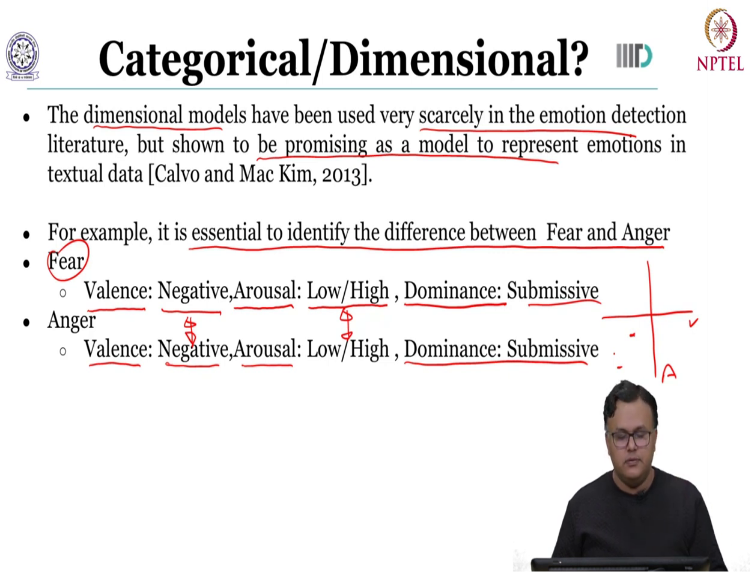
But the later it is more relevant in marketing or socio-political monitoring of the public sentiment, right. So, anger is more relevant in the marketing or socio-political monitoring.

But both fear and anger are negative; that means, you know we need to have a more fine grained representation of emotion when we are looking at text.

It is also shown that let us say when people are fearful, they tend to have more pessimistic view of the future. While angry people tend to have a more optimistic view, right. Even though both of these fear and anger are negative in their nature. But the intention for people that is different, right. So, when we are trying to understand the users their affective state through the text modality using fine grained annotation that is more clearer and more useful.

Further fear generally is a passive emotion while anger is more likely to lead to an action, right. So, there is very fundamental difference in what a user would be doing after they are you know experiencing these 2 emotions. Therefore you would like to have more fine grained transformation not just simply seeing positive or negative.

(Refer Slide Time: 23:12)



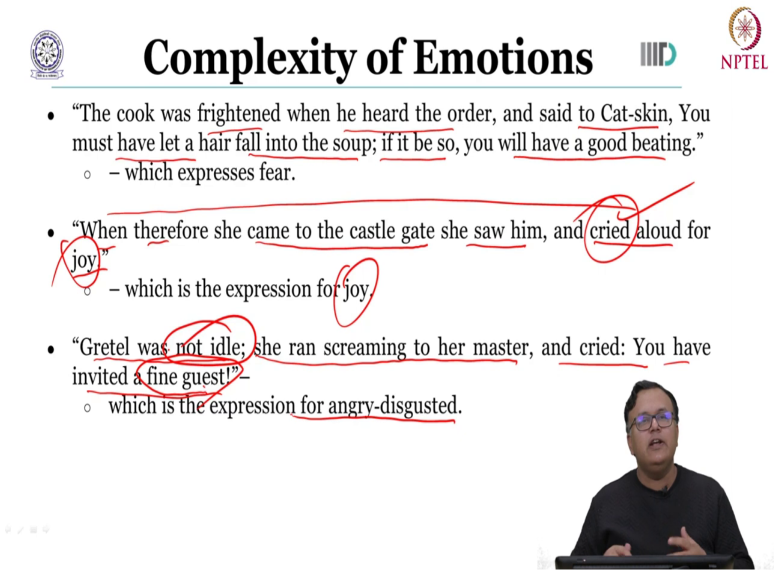
Now, the question of course, comes friends is when you are looking at text and you want to go fine grained emotion representation either we go for categorical classes or we look at the continuous emotion representations through the 4 dimensions. Now, the dimensional model that has been very less used in emotion detection literature. But of course, in the recent works it has shown to be more promising in the case of the text data.

Further, if you look at an example it is essential to identify the difference between fear and anger, right. You because of how fear and anger have different post objectives we will have to let us say look at the categorical data, but because it is difficult to label when can actually have the same represented on the continuous dimensions as well.

Now, when you look at fear with respect to its representation on the valence axis fear is negative and its representation on the arousal axis the intensity is low or high. And when you look at the dominance axis fear is submissive. Now, compare that with the anger emotion. On the valence axis it is negative.

So, this is same as fear on the arousal it could be low or high intensity and the dominance that is submissive. Now, the only difference would be that when you are actually having let us say the valence and arousal where are exactly these placed right, fear and anger. That would have an effect on better understanding of the emotional state of the user.

(Refer Slide Time: 25:24)



Now, in the same direction let us look at some examples, ok. These examples will convey to us how complex emotions can be when we are analyzing the emotion conveyed by a user in the text modality. So, let me read a statement friends. "The cook was frightened when he heard the order, and said to Cat-skin, You must have let a hair fall into the soup; if it be so, you will have a good beating".

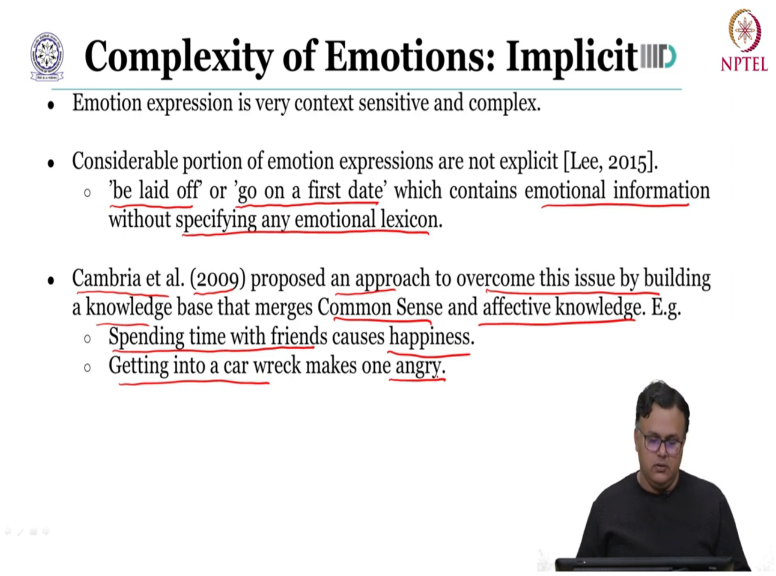
Now, here you have a complex statement where the sub parts are also reflecting different emotions. However, this mainly is expressing fear. Now, let us look at the second example the statement is "When therefore she came to the castle gate she saw him, and cried aloud for joy."

Now, if you notice the content, you have the word cried, you have the word joy is it sad is it happy. We can only understand the meaning in the terms of emotion when we look at the whole statement and the semantic meaning of it, right. So, this statement actually is an expression of joy even though the word cried is there in the statement, right. So, this actually reflects how complex emotion representation to text can be.

Now, let us look at another example friends. The statement is “Gretel was not idle; she ran screaming to her master, and cried: You have invited a fine guest”, ok. Now, if you look at the emotion this is actually an expression of angry and disgust. However, notice towards the end of the statement you have invited a fine guest. So, there is a bit of anger there is a bit of sarcasm as well, but it is only understood through when you analyze the whole statement, right.

So, you can actually see the relation between some words which are towards the end, but which are related to the words in the beginning, right. So, the emotion is conveyed when you start linking the words and this represents the complexity of emotions when you are looking at the text.

(Refer Slide Time: 27:58)

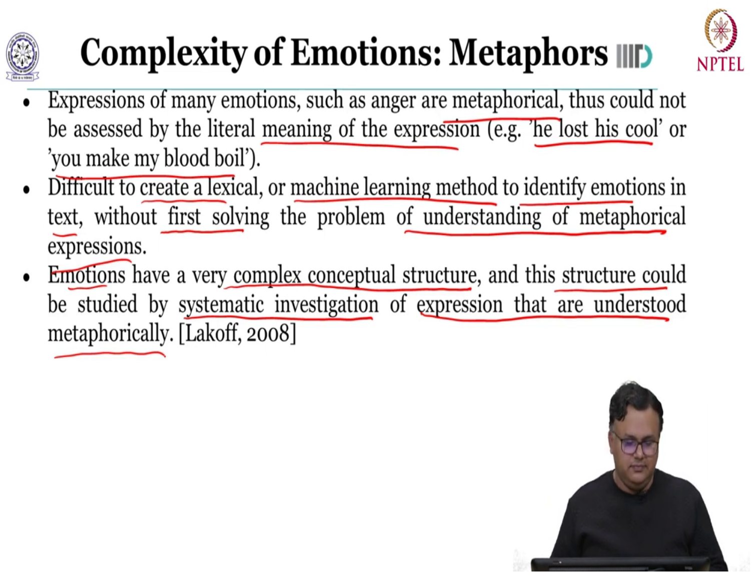


Now, emotions they can be implicit, right. Emotion expression is very context sensitive and complex. We have seen that in these three examples. It is noted that you know a considerable portion of emotion expressions they are not explicit. Implicitly through the text you would be actually trying to understand the conveyed emotion, ok. For example, 'be laid off' or another statement 'go on a first date'.

Now, these contain emotional information without specifying any emotional lexicon, right. So, there is an implicit emotion representation here. Now, in 2009 Erik Cambria and others they proposed an approach to overcome this issue by building a knowledge base which merges common sense and affective knowledge.

Example, spending time with friends causes happiness. Getting into a car wreck makes one angry, right. So, adding common sense and the affect knowledge, if this happens this would be the generic affect which would be there.

(Refer Slide Time: 29:26)



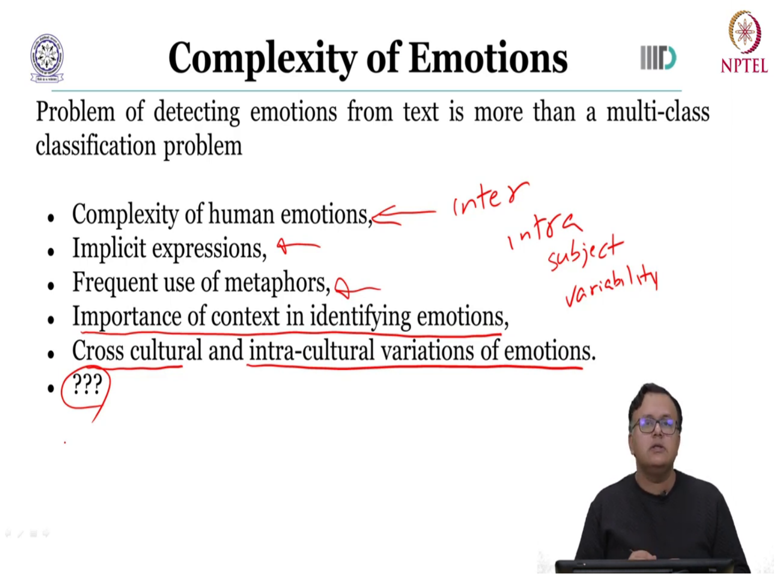
So, emotions can also be represented metaphorically, right. So, if you see expressions of many emotions such as anger, they are metaphorical. Therefore, they cannot be assessed by the literal meaning of the expression. Listen to this example friends, 'he lost his cool' or 'you make my blood boil'. So, these are metaphorical in nature. It is not actually boiling the blood it is actually the emotion which is conveyed anger, right.

Now, it is difficult to create a lexical or a machine learning method which can identify emotions in such text. And this is without first solving the problem understanding the metaphorical expression, right. So, in these kind of statements we need to understand the metaphor so as to be able to solve the redial of what is the emotion which is being conveyed by the text.

Now, emotions these are very complex conceptual structure and the structure could be studied by systematic investigation of expressions that are understood metaphorically. So, given that you know they are such a complex construct we need to understand the expression which are understood metaphorically.

So, that means, of course, you know if you wanted to create a system which could understand the implicitly presented emotions and the metaphorical ones you would need these datasets. And of course, in that you would need these samples where emotion is presented metaphorically.

(Refer Slide Time: 31:12)



Now, in the same direction friends the problem of detecting emotions from text is a more multi classification, multi class classification problem. This is simply because there is so much complexity of human emotion how people are presenting emotion what they speak, how they speak. Of course, you know this is all about the inter and intra subject variability.

Then some emotions or some speakers in a statement would speak and statement and then the emotion would be implicitly represented which saw some examples. Some subjects will very commonly use metaphors. So, that you know would mean that first we need to understand the metaphor itself.

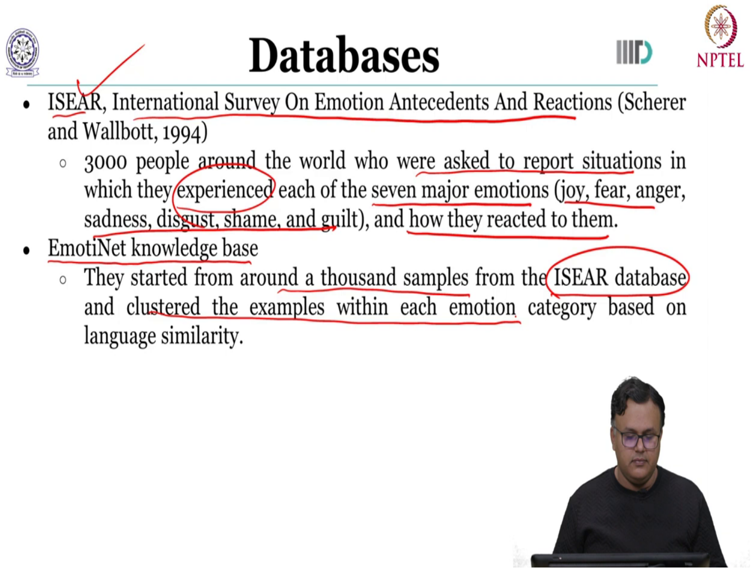
And then there is an importance of context in identifying emotions right, you need to understand in what context something was said, right. Is the context a comment about a product, is the context that someone is orating someone is telling about how their day was and in that particular use case scenario you would be then understanding the emotion, right.

So, the system needs to understand the context as well. Further as we have seen with voice based emotion and face based emotion analysis, cross cultural and intra cultural variation of emotion is there. Even if you are looking at the same language different people will be expressing the same context in different manners.

Same content in different manners which would mean that there would there is a lot of variation even though let us say the subjects are trying to convey the same emotion, but in different styles, right. So, this actually makes the task complex and that means, we will be required to have a multi class approach here.

And then of course, there are a lot more challenges right, when you are talking about text the moment you go to different languages. Their emotion will be represented differently, metaphors would be different, phrases would be different then the systems would need to adapt to these different language and cultural scenarios. So, there are a large number of challenges when we are looking at the text based emotion analysis.

(Refer Slide Time: 34:05)

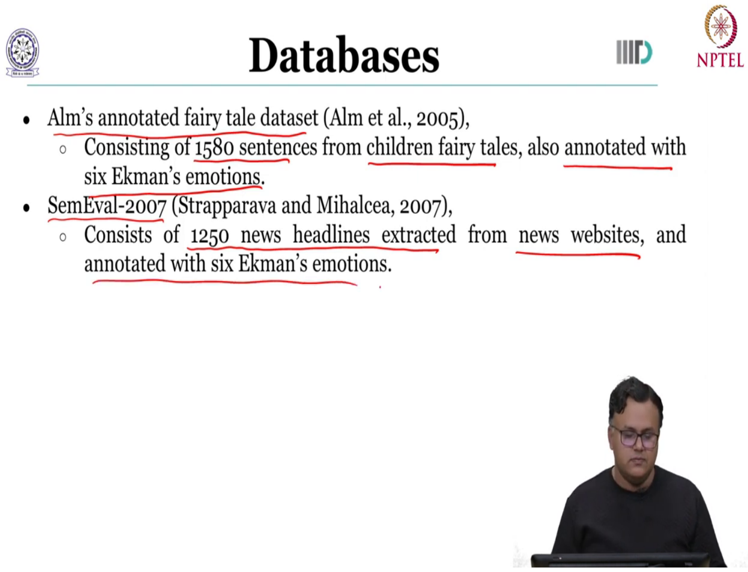


Now, friends let us look at some of the standard resources which are available for text based emotion and understanding. There is a dataset called ISEAR which is essentially the International Survey on Emotional Emotion Antecedents and Reactions by Scherer and Wallbott from 1994.

Now, in this there are 3000 people who were asked to report situations in which they experienced each of the seven major emotions. So, it is a categorical approach and they were also asked how they reacted to them, right. So, it is essentially that the user will recount about an event and then they would then you know write about it and when they are writing about, they are recalling an event then the emotion would be elicited.

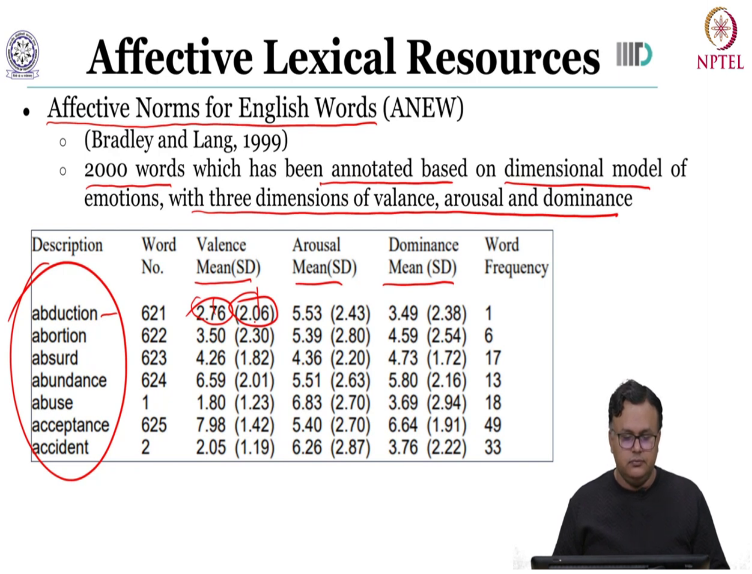
Now, the other one friends is the EmotiNet Knowledge Base in this the authors started from around a 1000 samples from this ISEAR dataset and clustered examples within each emotion category based on the language similarity.

(Refer Slide Time: 35:16)



Now, another dataset is alms Annotated Fairy Tale dataset proposed in 2005. In this dataset there are 1580 sentences from children fairy tales and they annotated with the categories from the Ekman's classical categorical emotion representation. Then there is the SemEval-2007 data by Strapparava and Mihalcea. In this we have 1250 new headlines extracted from news websites and these are again annotated for the categorical emotions.

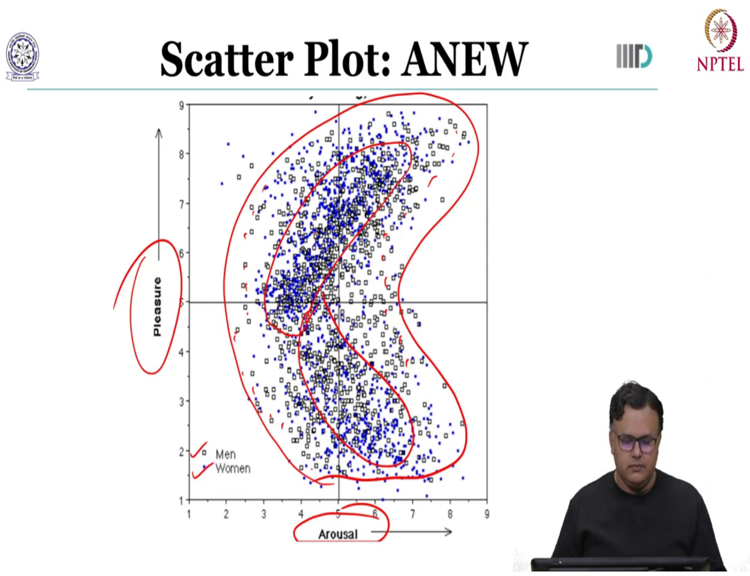
(Refer Slide Time: 35:55)



Now, let us look at another resource. Friends, this is also very commonly used resource this is called the Affective Norms for English Words ANEW. In this case the ANEW contains 2000 words which are annotated based on the dimensional model. So, these are about the valence and arousal with the dimensions of valence arousal and dominance.

So, here for example, you see the words and the mean valence arousal and dominance that is presented here for these particular words and in the bracket, you see the standard deviation. So, these are again collected from a large number of labelers and that is how you have the mean s and deviation.

(Refer Slide Time: 36:45)

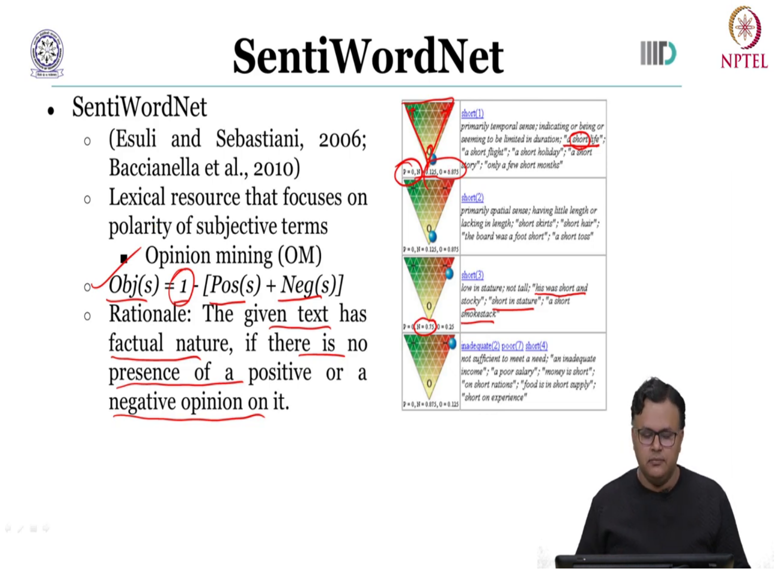


Now, if we observe the plot of pleasure versus arousal and the data points are for men and women, you notice here these are quite inter related they are overlaying. However, you will see the data coming from men labelers that is actually a bit more on the outside female data label.

So, essentially the arousal and pleasure dimensions they are more concentrate. This is just to show you that there is a bit of a difference when it comes to the labels as well based on the difference in gender. Of course, you know when you have n number of labelers who are labeling the app perceived affect of certain statements we are using text in this case then we are going to observe some difference.

And that is why we would like to compute basic statistics around the labels which are assigned by different labelers so that we can have a final label for the statement.

(Refer Slide Time: 37:51)

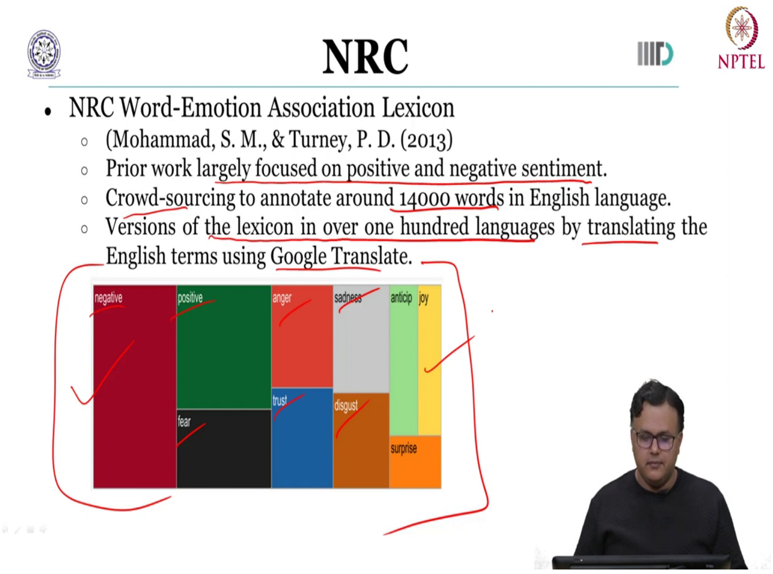


Another resource friends is the SentiWordNet. Now, in this case the lexical resources that focuses on the polarity of subjective term. So, this is a bit different approach. What we are saying is for mining of the opinion you can actually have an objective score which is 1 minus positive plus negative, ok. Now, the rational is as follows: the given text has factual nature. If there is no presence of a positive or a negative opinion on it, ok.

Now, let us look at an example in this case. So, here you see this is positive, this is negative and this is the opinion. So, if you pick up a statement for example, a short life. Now, in this case there is a word with a bit of negative connotation here, right. So, nothing positive P is equals to 0, negative is 0.125 and then of course, we are subtracting. So, you know this is the opinion.

Let us take another example friends. So, in this case you know let us say the statement is his was short and stocky, short in nature, a short smokestack, right. So, you have more negative connotation and the objective score that is 0.25 as compared to here where you had 0.875. So, this is the opinion about the statement.

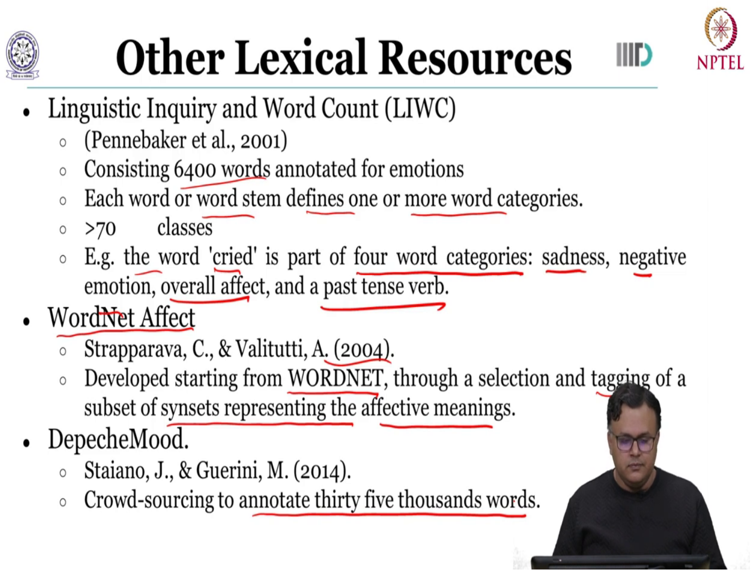
(Refer Slide Time: 39:35)



Now, let us look at another standard resource this is the NRC word emotion association Lexicon. Now, collected in 2013 this is based on prior work which largely focus on positive and negative sentiments, ok. So, what the authors they said well, let us crowd source and annotate 14000 words in English language. When you are crowd sourcing you can have a large number of labelers who are annotating the data.

And they have different versions of the lexicon in over 100 languages which are translated using the Google Translate and this is friends the heat map, ok. So, if you look at the sentiment this just shows the overall label. So, this has the highest number and of frequency and this has the lowest frequency.

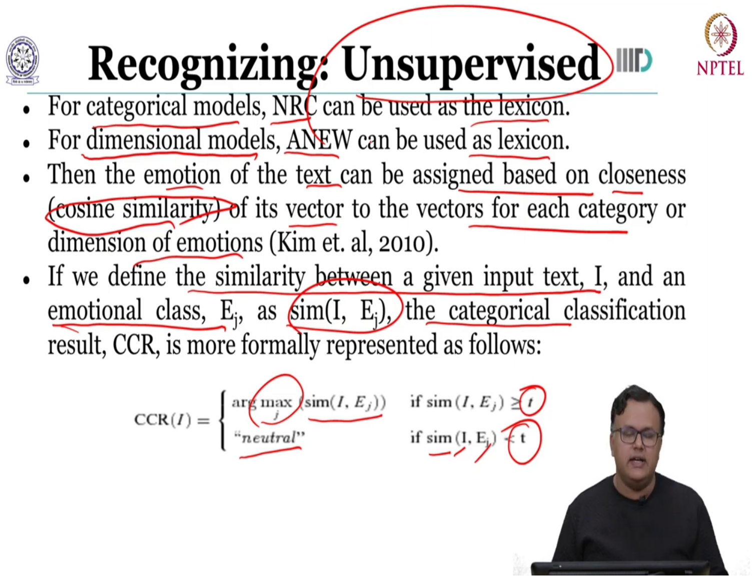
(Refer Slide Time: 40:30)



Now, some other lexicon resources this is the Linguistic Inquiry and Word Count LIWC another very commonly used resource it was proposed in 2001. Contains 6400 words which were annotated for emotions and each word or word stem defines one or more word categories. So, there are more than 70 classes. An example is the word ‘cried’ now it is part of four word categories sadness, negative emotion, overall affect, and a past tense verb.

Now, another one is friends WordNet Affect proposed in 2004. In this case it was developed from word net through a selection and tagging of a subset of synset representing the affect meaning which is subset of the work word net dataset. Then there is the DepecheMood which was proposed in 2014. And this one actually again is based on crowd-sourcing to annotate 35000 words. So, this is actually a bit larger resource as compared to the earlier ones.

(Refer Slide Time: 41:45)



Now, typically for categorical models, NRC data that is used as the lexicon. When we are talking about the continuous dimension the ANEW that is used as a lexicon. Now, further friends the emotion of the text can be assigned based on the closeness cosine similarity of the vectors to the vectors for each category. So, this is one way.

You are saying well, I have a representation for the text. Now, I am going to use the cosine similarity distance metric and I am going to assign the category or dimension. So, you would have some samples, some data points. Now these data points would have some category or some dimensional emotion based intensity assigned to them and ANEW sample comes in.

You compute the cosine similarity between this sample and the samples which already have the labels and you assign the one which is closest, right. So, the closest sample from the data will use those labels and assign it to the new sample.

Now, this simply means, well you know you define the similarity between a given input text I and an emotion class E of j and then the similarity for the categorical classification is formally represented as simply saying well, I want to compute the maximum for the similarity. If the similarity is greater than a threshold, if it is greater than a threshold you say well it is a non neutral emotion.

If it is neutral then essentially you are saying that the similarity between the input text and the emotion class that is less than a subtle threshold, right. So, in you can define the class in such a manner. But it simply means two things you need a representation for the text; you have the cosine similarity metric you can use some other metric for computing the distance as well.

And you keep a threshold where you say well, if the distance is larger than this particular threshold there is an emotion conveyed, if the distance is less than the threshold then this is a neutral statement. So, this way you can have unsupervised learning for the understanding of emotion using the text. So, friends with this we reach towards the end of lecture 1 of Text Based Emotion Analysis.

What we discussed was, why is text based analysis important, what are the applications, how the subtle details in topography, how fonts are represented, how the basic components inform the (Refer Time: 44:45) when they come together they would represent different emotions to the user and have an effect in inducing emotions in the user. And later on, we discussed about some of the most commonly used resources for text based affect understanding.

Thank you.