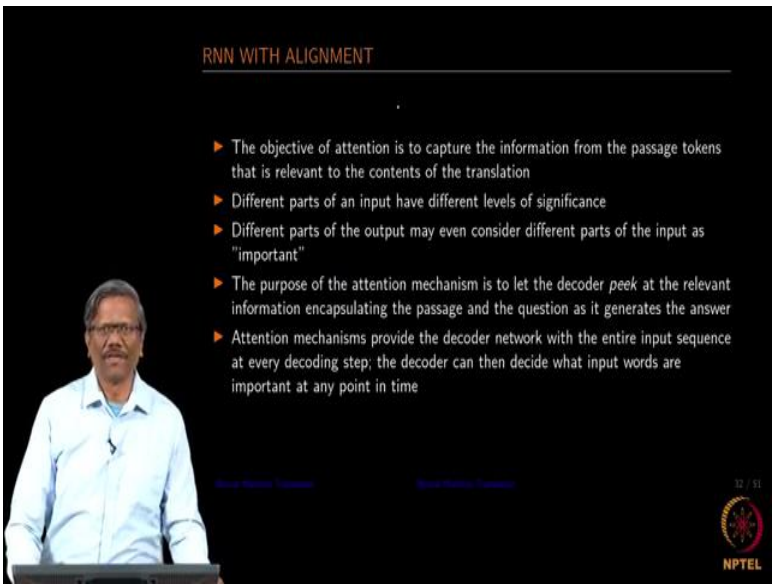
Applied Natural Language Processing
Prof. Ramaseshan Ramachandran
Department of Computer Science and Engineering
Chennai Mathematical Institute, Madras

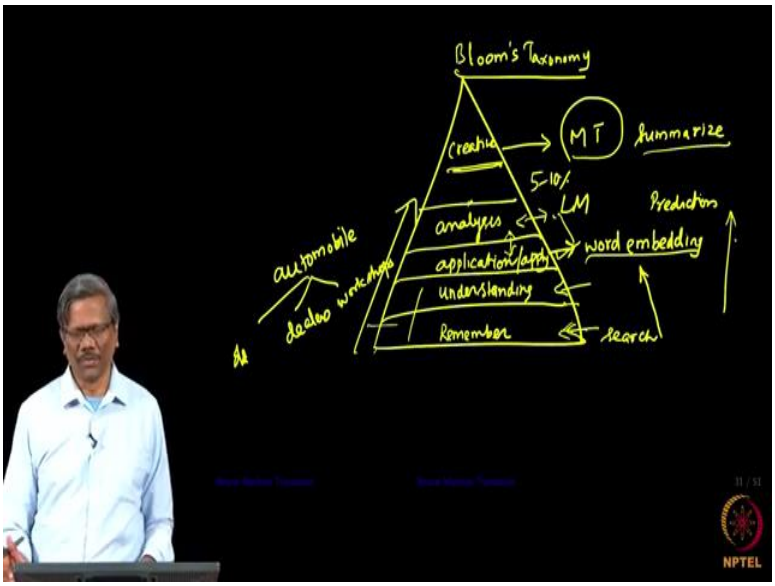
Lecture - 77
Recap and Connecting Bloom Taxonomy with Machine Learning

(Refer Slide Time: 00:16)



Ok. In the last segment we saw how we can translate from one language to the other using an encoder and a decoder both are RNNs, right. Have you noticed how we are really progressing? I have been talking about this every time you know when we take up a new topic, in natural language processing. It is very similar to what you probably would have noticed in Bloom's taxonomy. I spoke about this at the beginning of the course; in Bloom's taxonomy you know there are about 6 levels.

(Refer Slide Time: 00:57)



It would be like a pyramid and then in the pyramid what you will first have is remember. There is even you are thought something you remember that and then you would be able to recall that what you have been told, right; so, it is like the search operation that we have in the Bing search or Google search or if another search engine. Just given a word the system knows what it should find out based on the input word that we are given. So, we can equate that to our search operations, right.

So, without really understanding what is the meaning of the word that you are inputting, the search operation the basic search operation I am talking about in the early search engine days, which recovers the information related to what you have input as the query data right, but it is a simple retrieval operation, ok. Then next is about understanding. So, this is; so, from here we want to move to the next level of what you are really inputting, right. So, in this search it if it is a very blind search based on the keyword, you will not really like it; that is how it was at the beginning of these search engine days.

And, then later every word that you are inputting you know for example, if you say to use a word buy or use bought. So, both are taken to be as one word, even if this is not found in the text document. So, it will still bring a document that contains the word buy even if you had input bought into that, right. So, it tries to understand what you are really doing. So, some of the search engines started adopting that and then there came the understanding aspect of that, it tries to understand what is the kind of keywords that you

are inputting and then tries to find the equivalent of that. Even if the same keywords are not found in the document, it still retrieves, ok. So, then this is the next level trying to understand the query itself.

And, then the third one would be your application right, application or you call it as apply. Whatever knowledge that you have gained by reading through it, by understanding some of the syntaxes of that or the meaning of that you try to apply to a different scenario. So, this is where our word embedding started. So, we understand the query, we understand what the meaning of that word to some extent and then word embedding, what it does is even though it is not exactly finding out the meaning of the word that is found in the query; tries to figure out what are the neighborhood words that it would have seen in the earlier training days, right.

So for example, if you train the neural network with a huge collection of English documents, it tries to understand the meaning of the words; we have seen that in our word embedding classes and so on right. So, it tries to understand the neighborhood words, not exactly the meaning of the word, but it tries to understand what could be the possible next word if I give the first two words and so on. And, also it relates certain combinations of words and then gives you that; so, to give you an example. So, if you search for automobiles, word embedding based searches would provide details and later to dealers, workshops may be an old car, dealers and so on.

So, for anything related to that, you know it finds as its neighborhood that came as part of the word embedding, right. So, we can relate to some extent Bloom's taxonomy to the word embedding at the third level. And, then if you go to the next one which is the analysis part of that, I would say that these two are somewhat combined in the word embedding, ok. So, it tries to really analyze what should be the next for example, in language modeling, right. So, it read all the documents belonging to a corpus and then try to find out what could be the next word right, it is a prediction type.

So, for that it does a lot of analysis as part of the training process, right. So, we can relate the Bloom's taxonomy of analysis to both language modeling as well as word embedding there. And then the last one, I am going to be combining that as creative part, so instead of using evaluation and creation I am going to combine both as one ok. So, this is where I think we have the machine translation, the creative part is about writing your own

sentences, ideas, and then based on what you have learned in the process. So, machine translation we can closely say that it is trying to be very creative give when the system has learned the translations based on the parallel corpora. It is given a new sentence; it is now going to be creating a sentence for you which is a creative process, ok.

So, this is where machine translation and then summary, summarize a huge document into a few sentences. So, those aspects come into that, this is basically a creative aspect. So, we are slowly taking the machine from the basic remembering aspect to the creative aspect through machine learning as an algorithm, ok. So, this is a very small comparison that I thought we should make in the machine learning idea. So, this is a very simple comparison we probably should make, ok. So, by understanding Bloom's taxonomy level for example, now we can rate the machine also whether it is doing the job of just a recalling the values or it understands what you have provided.

And, then gives you the meaningful output for you or it is able to apply certain knowledge it has captured by going through the documents and then give you some results out of that, ok. So, some more level of intelligence we have given to the machine and then later in terms of language modeling when it tries to predict the next word based on what it has seen earlier. It is similar to the analysis aspect may be a small or thin boundary between the creative and the analysis side, but the creative is very clear with respect to machine translation and summary. So, it is able to do something which humans were doing earlier, right.

So, it is a very creative process and then it also creates a lot of possible translation and many of them may not really be a good one, but out of which you know you can pick and choose some of the good translation that is coming out of it. So, the translation; the machine translation part that we have been talking about has reached a stage that may not be very creative like a human. But, at least it is able to replicate about 5 to 10 percent rather 5 to 10 percent of humans in this aspect. So, this is one thing that I thought you should know, right. And, for your information Bloom's taxonomy, you used in the many of the examinations, where they try to provide questions based on the pyramid.

You know there will be some questions in terms of identifying your creativity. There will be some questions that will try to understand how are you analyzing certain elements based on the data, that is given to you and then how do you really apply the knowledge

that you have gathered. So, there will be some questions on that and then there will be some recall questions that probably you would know very well, right. So, many of the exams follow this pattern. So, it is I thought it is a good idea to compare how machines also learn and then where we are really taking the machines to how with respect to this.