but for the emphasis on heavy evaluation, starting a trend that became a major feature in 1990s (Young and Chase, 1998; Sundheim and Chinchor,1993) [131, 157]. Work on user modeling (Wahlster and Kobsa, 1989) [142] was one strand in a research paper. Cohen et al. (2002) [28] had put forwarded a first approximation of a compositional theory of tune interpretation, together with phonological assumptions on which it is based and the evidence from which they have drawn their proposals. At the same time, McKeown (1985) [85] demonstrated that rhetorical schemas could be used for producing both linguistically coherent and communicatively effective text. Some research in NLP marked important topics for future like word sense disambiguation (Small et al., 1988) [126] and probabilistic networks, statistically colored NLP, the work on the lexicon, also pointed in this direction. Statistical language processing was a major thing in 90s (Manning and Schuetze,1999) [75], because this not only involves data analysts. Information extraction and automatic summarizing (Mani and Maybury,1999) [74] was also a point of focus. Next, we present a walkthrough of the developments from the early 2000.

3.1 A walkthrough of recent developments in NLP

The main objectives of NLP include interpretation, analysis, and manipulation of natural language data for the intended purpose with the use of various algorithms, tools, and methods. However, there are many challenges involved which may depend upon the natural language data under consideration, and so makes it difficult to achieve all the objectives with a single approach. Therefore, the development of different tools and methods in the field of NLP and relevant areas of studies have received much attention from several researchers in the recent past. The developments can be seen in the Fig. 3:

In early 2000, *neural language modeling* in which the probability of occurring of next word (token) is determined given *n* previous words. Bendigo et al. [12] proposed the concept of feed forward neural network and lookup table which represents the n previous words in sequence. Collobert et al. [29] proposed the application of *multitask learning* in the field of NLP, where two convolutional models with max pooling were used to perform parts-of-speech and named entity recognition tagging. Mikolov et.al. [87] proposed a *word embedding* process where the dense vector representation of text was addressed. They also report the challenges faced by traditional sparse bag-of-words representation. After the advancement of word embedding, neural networks were introduced in the field of NLP where variable length input is taken for further processing. Sutskever et al. [132] proposed a general framework for *sequence-to-sequence* mapping where encoder and decoder networks are used to map from sequence to vector and vector to sequence respectively. In fact, the use of *neural networks* have played a very important role in NLP. One can observe from the existing literature that enough use of neural networks was not there in the early 2000s but till the year 2013enough discussion had happened about the use of neural networks in the field of NLP which transformed many things and further paved the way to implement various neural networks in NLP. Earlier the use of *Convolutional neural networks* **(**CNN**)** contributed to the field of image classification and analyzing visual imagery for further analysis. Later the use of CNNs can be observed in tackling problems associated with NLP tasks like Sentence Classification [127], Sentiment Analysis [135], Text Classification [118], Text Summarization [158], Machine Translation [70] and Answer Relations [150]**.** An article by Newatia (2019) [93] illustrates the general architecture behind any CNN model, and how it can be used in the context of NLP. One can also refer to the work of Wang and Gang [145] for the applications of CNN in NLP. Further