NLU helps the machine to understand the context, semantic, syntax, intent, and sentiment of the text using various rules and techniques. NLG generates structured human-understandable data at high speed. NLU and NLG combines to form NLP.

Verb-based approach is used to determine the verbs in a sentence and to bind the subjects and objects.

Brill’s Tagger is used to tag the different parts of speech in a sentence.

Heuristic approach is used for text summarization.

Score based clustering algorithm is used for parsing text.

Segmentation using the Multilayer Perceptron model is used for obtaining textual information from pictures.

Tag-switching, inter-sentential and intrasentential.

Hough-transform is used for fuzzy feature extraction.

A hybrid model based on gradient features and coefficients of the wavelet transform can be used to recognize Bangla characters on a printed form.

CNN-based models with a Histogram of Oriented Gradient (HOG) feature can be used in detecting Bangla numerical digit recognition.

The procedure of developing a Bangla machine translation dictionary is as follows: determining contents of the dictionaries, determining the detail of information, determining the organization of the dictionaries.

Dictionary-based, rule-based, and statistical-based approaches

A RNN based network can be good for this purpose. Graded Recurrent Unit can be used for vanishing gradient problem of RNN. Neural networks consists of the forward pass and backward pass. The forward pass computes a weighted sum of their inputs from the previous layer and pass the result through a non-linear function. The backward pass is to compute the gradient of an objective function with respect to the weights of a multilayer stack of modules via the chain rule of derivatives.

Kimmo Koskenniemi’s two-level morphology technique uses finite-state morphological parsing which has 3 components for parsing system: lexicon and morphotactic, morphophonology and word-grammar component.

CFG can be used for generating a parse tree. Different CFG can be created for different indicative forms of sentences to make a parser.

The suffix stripping algorithm checks if a word has suffixes from the list and the cluster is assigned assuming the root word. Minimum edit distance resolves the stemming problem for inflected verb words.

The methodology of the POS layer tagging starts with morphological analysis of the words. In the beginning noun, analysis and verb analysis is done. Then the suffixes are classified based on number, postposition and classifier information. Verbs are classified into 6 paradigms based on the morphosyntactic alternation of the root. The suffixes are further analysed for person and honorific information. Then the ambiguity between a cardinal and a noun is resolved by a rule.

A tag vector is a sixteen-bit tag vector where parts of speech, person, mode, tense number and emotion are put in different lengths.

Collection of corpus and dataset

Search for implementation of Baum-Welch algorithm  
 Performing training  
 Final test against the gold standard for accuracy

A Hash table is used that maps the keys to values for storing data. The algorithm which is proposed works on three hash tables. The word list words are checked to examine whether it matches any root word stored in the hash map. If the word from the suffix list matches the root word, then this word is added in hashmap1, hashmap-2, and hashmap-3. If it matches multiple times, then the longest match is considered. Then the stored words in the hash maps are analyzed to figure out the potential words suitable for POS tagging words. In this case, the occurrences of the same tag were eliminated. Finally, the POS tag dictionary is generated using hash maps using the potential candidates.

A machine learning approach using stochastic gradient descent can be used to classify the Bangla question-answer. Here two-layer taxonomy, which has six coarse classes: abbreviation, entity, description, human, location, numeric, and 50 finer classes is used.

Possible pre-processing steps: Removal of emoticons, removal of proper nouns, and manual validation (by native speakers).   
2 methodologies for sentiment analysis: LSTM and neural network.

Multinomial naive Bayes can be used to detect spam from malicious Bangla text. Here punctuation marks, numerical values, and emoticons are extracted, and a TF-IDF vectorizer is used to extract the features.

Lexical features, syntactic features, semantic features, metadata and punctuation can be used to classify Bangla news as fake.

When we normalize text, we attempt to reduce its randomness, bringing it closer to a predefined “standard”. This helps us to reduce the amount of different information that the computer has to deal with, and therefore improves efficiency.

The extraction approach uses sentence scoring to evaluate the key sentences to focus on summarization. The process starts from pre-processing. The prime sentences are then identified by word analysis and sentence analysis based on length, distance, values, etc. Finally, the prime sentences are evaluated using sentence scoring methods. The final process includes aggregate similarities, final gist analysis, and sentence ranking.

Using apriori algorithm, one can determine the meaning of a Bangla word with multiple meanings.

To implement a naïve Bayes method, first we need to normalize the text. Then we need to convert it to Unicode format, remove the non-functional words and mark the unambiguous words. Finally we need to generate reference output files with the help of a Bangla dictionary.   
To implement a KNN-based algorithm, first it is needed to remove all the stop words. Then we can use a Bengali stemmer on the entire dataset.

Speech Signal => Speech signal Preprocessing => Feature extraction => Phonetic unit recognition => Language Modeling => Decoded Message

The speech signal processing part consists of speech starting and endpoint detection, windowing, filtering, calculating the LPC and cepstral coefficients, and finally constructing the codebook by vector quantization. The speech pattern recognition part consists of recognizing patterns using an ANN.

The acoustic model’s training process requires a Bangla phonetic dictionary, a language model, and Bangla acoustic data. The Google Translate application can be used to generate the phonetic transcription of Bangla words. The proposed Bangla speech recognizer model achieved 86.7% word recognition accuracy.

RNN can be used in BNLP for augmented noisy and general public speech recognition, speech processing and recognition, information extraction, named entity recognition, sentiment analysis etc.

DNN can be used in the preprocessing steps. For named entity recognition, DNN can be used to compare the base model. DNN can also be used to compare the CRFs model’s performance, whether the modern machine learning algorithm exploits the results more or not.

CNN takes the 2D structured MFCC features as input. The proposed system overcomes the drawbacks which researcher face while working with 1D structured input. And thus CNN can be used for obtaining noise-free numeral speech recognition.

The speech signal processing part consists of speech starting and endpoint detection, windowing, filtering, calculating the LPC and cepstral coefficients, and finally constructing the codebook by vector quantization. The speech pattern recognition part consists of recognizing patterns using an ANN.

Dynamic Time Warping (DTW), Hidden Markov Model (HMM), Linear Predictive Coding (LPC).

CRF means Conditional Random Field. It is a class of discriminative models.

Transfer learning is a machine learning method where a model built for a specific task is reused as a starting point for a model on another task. It’s re-using pre-trained models on newer problems.

The basic concept of ensemble learning is to train multiple base learners as ensemble members and combine their predictions into a single output that should have better performance on average than any other ensemble member with uncorrelated error on the target data sets.

Random Forest is a powerful and versatile supervised machine learning algorithm that grows and combines multiple decision trees to create a “forest.”

The authors of the paper had very few words of themselves on the paper. Though they discussed the methodologies and performance metric of many other papers, they did not give their own opinions or suggestions regarding those mentioned papers. I think that’s where the authors could improve more.

B. Kundu and S. Chandra, “Automatic detection of English words in Benglish text: A statistical approach,” in 2012 4th International Conference on Intelligent Human Computer Interaction (IHCI), pp. 1–4, IEEE, 2012.

<https://books.google.com.bd/books?hl=en&lr=&id=KGIbfiiP1i4C&oi=fnd&pg=PR5&dq=natural+language+processing&ots=Y4Bfy0HFO6&sig=Kysv2OU0gpLakJLP4hzPfcz69xQ&redir_esc=y#v=onepage&q=natural%20language%20processing&f=false>

<https://link.springer.com/chapter/10.1007/978-3-030-18305-9_9>  
<https://link.springer.com/chapter/10.1007/978-3-030-32236-6_48>  
<https://www.sciencedirect.com/science/article/pii/B9780128235195000014>  
<https://ieeexplore.ieee.org/abstract/document/8752407>  
<https://ojs.aaai.org/index.php/AAAI/article/view/5108>