**Introduction:**

**1.1 Transliteration**

A transliterated form of language is defined as the form in which alphabet of one language is used for representing another language. For example, “Ami bari jabo” is a Bangla transliterated sentence using the Roman alphabet. The use of transliteration is huge in many languages. The roman alphabet is used in western and eastern European languages extensively. The roman alphabet is also used in Turkey, Vietnam, Somalia and in east Africa(Swahili Language).

**1.2 Transliteration in Bangla Language**  
  
 We use the transliterated form to write Bangla in various sectors. The maximum use of the transliterated Bangla sentences is in social media, like Facebook, Instagram and twitter. Another sector where this transliterated form is used extensively is in chatting applications, like Messenger, WhatsApp, Viber etc. We also use this form to write in online blogs and web portals.

**1.3 Importance of Identification of Interrogative Sentences from Transliterated Bangla Sentences:**  
  
 In all these sectors mentioned previously, we are generating considerable number of interrogative sentences every day. The identification of interrogative sentences is very much important for various reasons. From big data analysis to user behavior study, this identification can play a significant role. Service providers can study the client behavior, expectations, demands, queries from the interrogative sentences posted by the clients in the form of Facebook status, comments, tweets, online blog posts, live chats etc. This identification is also important for the development of smart assist applications, medical applications, question-answering based applications, user interactive applications and so on.

**1.4: Bangla Interrogative Sentences in Transliterated Form:**  
  
In transliterated Bangla, the interrogative sentences come in various forms. In many cases these sentences differ to great extent from the traditional grammatical format of the interrogative sentences. Many people do not use question mark at the end of the interrogative sentences. This common practice makes the challenge of identification very hard. According to [1] ,30% of the online questions do not contain any question mark at the end. One important thing is that the question mark at the end of the sentence does not necessarily express questioning property all the time. For instance, “Tomar eto boro spordha?”. This sentence expresses exclamation rather than questioning property. Another common thing is that people often omit the Bangla question words (ke, kokhon, kar etc.) while writing interrogative sentences. For instances, “Apni jacchen tahole?”, “Apni shotti chole jaben?” etc. There are some sentences where the presence of question word does not indicate the questioning property rather the question words may act as a linker. Figure 1 illustrates an example of such sentences taken from our dataset.

**1.5 Data Collection:**

We have collected our dataset from cricket based Facebook public groups where a huge number of questions (as comments) are generated by the members in the transliterated form. The variation of the dataset is remarkable. The dataset contains all types of Bangla sentences. This dataset is used as our training and experimenting corpus. We have also included a dataset regarding SUST admission test, which we have collected for our another project. This dataset also contains a great number of Bangla interrogative sentences in the transliterated form.

**1.6 Approach of Identification:**   
In this paper, for the identification of interrogative sentences, we have followed both the rule based approach, Supervised learning approach and the deep learning approach. In the rule based approach we have followed the rules of grammar and some rules which we have made by observing and analyzing the dataset. The drawback of this approach is that due to variations in sentences no set of rules can identify the interrogative sentences accurately and efficiently. For better accuracy and efficiency, we have used supervised learning. In supervised learning, We have used the State Vector Machine(SVM), Logistic Regression, Multilayer Perceptron(MLP) and k-Nearest Neighbor(KNN) algorithm for the classification and identification. We have used Convolutional Neural Network(CNN) for the deep learning approach.

1. Example of Bangla Transliterated sentences where question words are used as linker

1.Ami kokhon bari jabo ta ekhono Janina

2.Ami ki kori tate karo kichu jay ashena

3.Tini janena ki karone tara asheni

4. Khela koto tarikhe khela ache ta ami jani na

5.Mather kothay she darano ache ta dekha jacche na.

6.Brishti shulu hole kivabe khela shuru hobe ta ekhon bola jacchena