To test variety of patterns and NLP algorithms and their combinations for any task given. Therefore, for comparison tool must allow us to add external plug-ins into the framework. Consequently, system needs to have the supportive data types and structures that allow for the scripting of the recognition tasks for all potential processing. Earlier MARF framework was stand-alone, sequential and had limited support for multithreading. After learning it evolved and made distributed but it was manual so again improvised and made autonomic[MN].

To identify the different features of speaker like their gender, their accent etc. It uses concept of mean clusters and median clusters to find out the best possible combinations of algorithms[MH].

Scripting in required application is difficult task by providing context of all parameters. Therefore we need to make the syntax simpler which overload context operators to accept various types of arguments and return types @ and # which are taken from Generic Intensional Programming Language (GIPL) which in turn helps to achieve this task[MS].

MARF provides great usefulness to researchers to decide different combinations of algorithm. It also provides facility to choose best suited algorithm combination for each task. The estimate of algorithm combination is based on statistical estimators and NLP parsing and many other modules[MK].

NLP techniques are used to analysis of source code with safety to find vulnerabilities and weakness in code. For that we used MARF’s NLP framework and MARCRAFT application. In these experiments unigram alone was used because it has produced good precision and they are the fastest among all other but signal pipelines[MT].

MARFCAT is a MARF-based code analysis tool which is presented at the Static Analysis tool exposition (SATE) workshop 2010 and collocated with the Software Assurance forum.The methodology behind static source code analysis contains two core principles: Machine learning and Spectral and NLP techniques which use signal processing techniques[MC].

To identify speaker’s gender and accent through Machine Learning. MARF uses SpeakerIdentApp as a testbed which can be used as a tool for comparing different algorithms as well as it allows dynamic module selection based on available configuration options[MA].

The writer identification techniques skeletonizing, contouring, line-based and angle-based feature extraction are highly accurate but its time consuming for large volume of digital data of handwritten material. Therefore, by modifying MARF’s Pipeline, WriterIdentApp and Resolution we simulate “quick visual identification” of the hand writing of the writer[MM].

References:

[MT]

[MA]

[MM]

[MN]

[MS]

[MC]

[MK]

[MH]