ATTEMPTING AN INTEGRATED NN-GA (NEURAL NETWORK GENETIC APPROACH) MODEL ON SRS TO ENHANCE THE ACCURACY AND PRECISION FOR EXTRACTING FUNCTIONAL REQUIREMENT

Himanshu Dahiya

Bachelor of Technology (B.Tech), IT Manipal University, Jaipur-303007(Rajasthan), India

ABSTRACT

The suggested method for extracting Functional requirement from the SRS by using concept of clustering with neural genetic approach. Extracting functional requirement is a different task for judgment, a pattern in the functional requirement that is communicated as regular expression in the contribution of SRS. We proposed and implement pre-processing of SRS by text mining and classified by cluster based supervised leaning. We reduce the complexity and increase the accuracy functional requirement extraction by using text from SRS and then clustered then classified by cluster based supervised learning. It targets at providing an automatic supervised Functional requirement content mining based on mode based structure that divides the SRS documents when a shared pattern is found. We could detect and remove local noises that rendered the Functional requirement by non-functional requirement into well-formed SRS documents that enhanced the extraction process. We analysed the proposed approach by precision, recall, accuracy and F1 measure.

I. INTRODUCTION

"Intelligence is the ability to adapt to change"- Stephen Hawking. This statement appropriately features the significance of adjusting to evolving condition. Each product experiences certain exercises under programming improvement life cycle, and Agile strategies have risen as one of the most prevalent methods for programming advancement as of late because of their capacity to adjust to changing necessities and evolving condition; in this manner, making these techniques able for a present-day business situation as per tenth yearly coordinated overview. Be that as it may, uniform client association all through the undertaking, delegating suitable group to adjust to nimble approach, positioning of changes to be obliged in programming, looking after straightforwardness, trouble in scaling spry systems to bigger tasks and settling on the agreement terms represent the significant hindrances engaged with the light-footed improvement. Large systems involve large separate teams working on separate interacting systems may be at different locations. Large systems also have external development restrictions and rules which limit their extent of flexibility. Also, managing large systems and involving various stakeholders of large projects is different. Adapting agile development for larger systems involves scaling up process which needs attention towards design documentation, arranging communication and meetings among different teams and careful maintenance of continuous integration of builds. Scaling out is required to introduce agile methodology in large organizations (following traditional approaches) and making agile methods

compatible with them. As compiled by agile approach provides benefits to both development team as well as the client by addressing project drawbacks like schedule predictability, scope creep and budget control etc. AI is essentially a subfield of man-made brainpower which loans PCs the learning capacity without unequivocally programming them. This includes improvement of those PC programs which have the capacity of showing themselves for developing and changing when they are presented to new information. Python is an abnormal state, translated, intelligent and object-situated scripting language. Python is intended to be profoundly comprehensible. It utilizes English catchphrases habitually whereas different dialects use accentuation, and it has less linguistic developments than different dialects.

II. METHODOLOGY

Step 1: Archives containing different prerequisites particulars (practical just as non-useful necessities) are gathered from different sources.

Step 2: These records are pre-handled to create the required content for PC examination. This pre-handling through content digging readies the content for PC justifiable portrayal and concentrates the fundamental, helpful issue from the content. This is done in three stages: Tokenization: It changes over a flood of characters into tokens (by and large word tokens). Delimiters like space, accentuations, and so forth are utilized for isolating single word token from another. Words in the word stem are decreased to their root structure like 'composes', 'composing', 'composed', 'composed' these all relate to a solitary root "compose".

Step 3: When the content is prepared, the record term weight is spoken to as Term Frequency-Inverse Document Frequency which assesses the significance of an archive word in a gathering of reports. Terms that catch the report embodiment are progressively visit and have high Term Frequency though great terms that segregate a record from others happen less in the archives and have high Inverse Document Frequency. Dab item is taken of both to figure a last Term Frequency-Inverse Document Frequency.

Step 4: Clustering is utilized to segment the information into sets of comparable things. Every one of the sets is known as a bunch. This procedure targets expanding union in a solitary bunch and limiting coupling between at least two groups that is, attempting to diminish the intracluster separation and increment between bunch separations. The grouping system utilized in the technique is a hybridization of a level and hard bunching calculation K-implies with an agglomerative base up various levelled grouping strategy UPGMA (unweighted pair group method with arithmetic mean). Effortlessness and productivity of K-implies is utilized to create an enormous number of bunches and afterward UPGMA is connected to refine these groups with the goal that their quality is improved.

Step 5: Supervised learning is connected to execute two classifier models: where, True Positive is effectively anticipated positive, False Positive is erroneously anticipated positive, True negative is accurately anticipated negative and False Negative is dishonestly anticipated negative. Pseudo Code of Proposed Algorithm Following is the pseudo code of the proposed calculation:

Step 6: Results from the above two classifier models are then analysed and compared based on Precision, Recall, and Accuracy which are calculated by the Formulae mentioned below:

$$\text{Precision} = \frac{\sum \text{True Positive}}{\sum \text{Predicted Positive}} = \frac{\sum \text{True Positive}}{\sum \text{True Positive} + \sum \text{False Positive}}$$

$$\text{Recall} = \frac{\sum \text{True Positive}}{\sum \text{Actual Positive}} = \frac{\sum \text{True Positive}}{\sum \text{True Positive} + \sum \text{False Negative}}$$

$$\text{Accuracy} = \frac{\text{Correctly predicted observations}}{\text{Total number of events}}$$

$$= \frac{\sum \text{True Positive} + \sum \text{True Negative}}{\sum \text{True Positive} + \sum \text{False Positive} + \sum \text{True Negative} + \sum \text{False Negative}}$$

Step 1: Input the data with features and labels.

Step 2: Apply the neural network

Step 3: Training neural network

Step 4: Evaluation of Mean Squared Error

Step 5: Start genetic algorithm t=0

Initialize the random population

Step 6: Increase counter t+1;

Step 7: Evaluate

Step 8: If weight is optimized, stop genetic algorithm else go to step6.

Classifier	Precision	Recall	Accuracy
SVM-RBF	86.86	86.97	85.25
NN-GA	87	87.23	90.23

SVM with RBF kernel: SVM can just perform direct characterization. RBF portion is joined with SVM to give the ability of non-straight arrangement. This enables the calculation to fit the most extreme edge hyper plane in a changed element space. Neural system with hereditary calculation: Use of the hereditary calculation diminishes the mistake of neural system.

III RESULTS

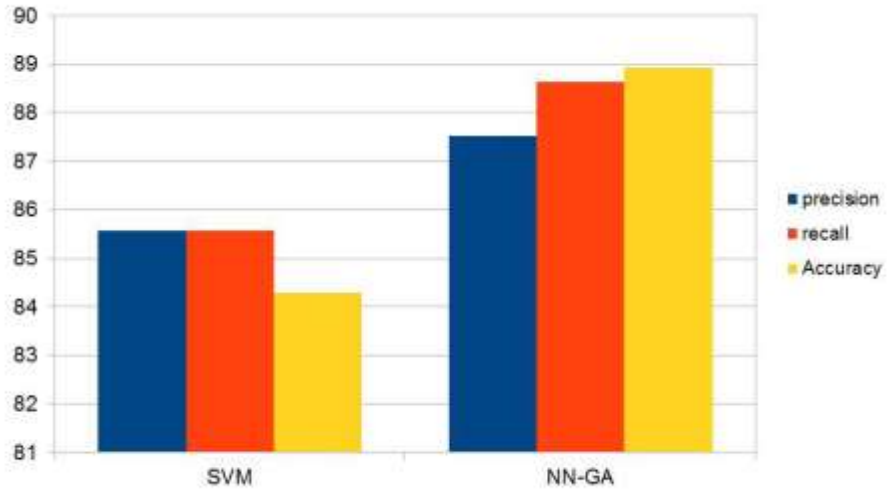

Table 4.1Comparative Analysis for EU eProcurement data set:

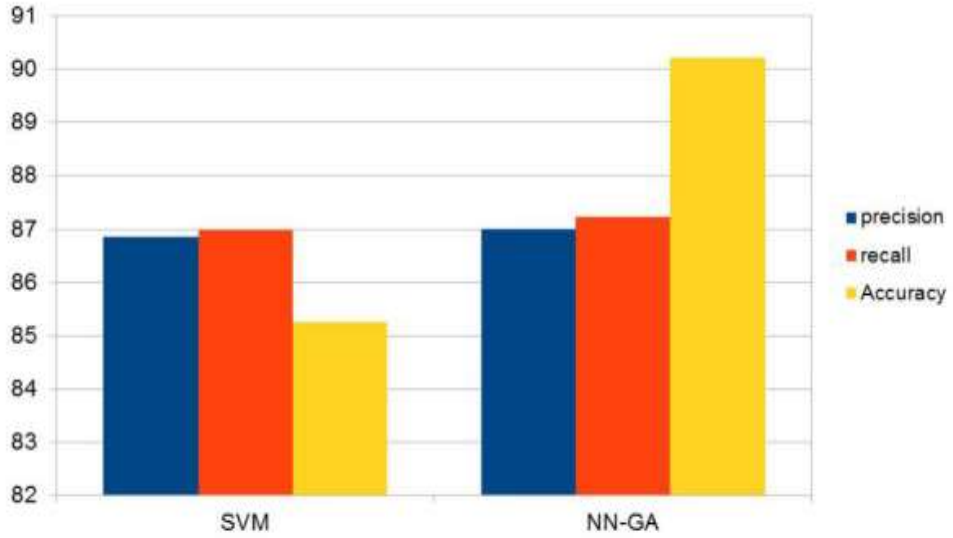

Table 4.2 Comparative Analysis for PROMISE data set

IV. CONCLUSION

Agile development has risen out as the most prevalent methods for programming advancement as of late. Lithe offers a wide range of systems every one of which has its own arrangement of advantages and disadvantages and might be chosen by programming organizations dependent on their condition, venture types, accessible assets, and different imperatives. Necessity building is a pivotal piece of this improved method. This paper proposes a programmed methodology dependent on regulated learning for an arrangement of useful and non-utilitarian necessities in light-footed advancement. The proposed technique does as such by utilizing the idea of grouping with the neuro-hereditary methodology. The necessities in prerequisite building venture of the procedure are imparted through ordinary articulations through Software Requirement Specification (SRS).

Classifier	Precision	Recall	Accuracy
SVM-RBF	85.56	85.56	84.29
NN-GA	87.52	88.62	88.92