SN Computer Science (2023) 4:332
https://doi.org/10.1007/s42979-023-01786-1

ORIGINAL RESEARCH

Affirmative and Negative Sentence Detection in the Brain Using SVM-RFE and Rotation Forest: An fMRI Study

Ashish Ranjan[1] · **Vibhav Prakash Singh**[2]

Received: 15 February 2023 / Accepted: 14 March 2023 / Published online: 15 April 2023

Abstract

Studies on the brain's response to positive and negative stimuli using non-invasive techniques such as fMRI and PET scans have been widely researched. The use of fMRI in particular, to study sentence polarity detection in the brain, has become a captivating field. This paper analyzed fMRI data of affirmative and negative sentence processing and developed a model using a correlation-based subset evaluator and NBTree classification to classify sentence polarity with an average accuracy of 92.9%. Further, the accuracy of the result is enhanced up to 100% using SVM-RFE and Rotation Forest. Our analysis of selected voxel sets in the brain showed that the calcarine sulcus and right and left dorsolateral prefrontal cortices play a significant role in determining sentence polarity, while areas such as the right posterior pre-central sulcus, right supramarginal gyrus, and right frontal eye fields have a limited contribution.

Keywords fMRI · Voxel · Sentence polarity · Greedy stepwise · NBTree · Decision tree · Naïve Bayes · SVM-RFE · Rotation Forest

Introduction

Functional magnetic resonance imaging (fMRI) works on the principle of blood oxygen level-dependent (BOLD) contrast. When an area of the brain is active, there is an increased demand for oxygen, leading to increased blood flow to that area. This increased blood flow causes a change in the magnetic properties of hemoglobin in the blood, which can be detected by the fMRI machine. By measuring these changes, fMRI can create images that show which parts of the brain are more active during specific tasks, allowing researchers to infer which parts of the brain are involved in specific functions. fMRI studies can be used to investigate the neural basis of sentence polarity detection, which is the process of determining whether a sentence has a positive or negative sentiment. During an fMRI study, participants may be presented with a series of sentences while lying in the fMRI machine, and their brain activity is monitored as they evaluate the sentiment of each sentence. The resulting fMRI images can reveal which brain regions are involved in sentence polarity detection, and how different regions interact to perform this task. This information can provide insight into how the brain processes language and emotions and can inform the development of computational models for sentiment analysis. Understanding the state of the brain has always intrigued people. It is now possible to record the brain's activity while a person is doing particular tasks thanks to the development of new non-invasive techniques such as fMRI and PET scans. Using magnetic fields, fMRI creates a three-dimensional image of the brain and demonstrates how various brain regions respond to a task [1]. The brain's patterns of activation and the specific regions that are active during particular activities, such as viewing images or processing words and sentences, have been identified using this data.

Neuroimaging data have helped advance the study of how the brain processes language, particularly syntax (sentence

This article is part of the topical collection "Machine Intelligence and Smart Systems guest" edited by Manish Gupta and Shikha Agrawal.

✉ Vibhav Prakash Singh
vibhav@mnnit.ac.in

Ashish Ranjan
ashish.ranjan@vitbhopal.ac.in

[1] VIT Bhopal University, Sehore, India

[2] Department of Computer Science and Engineering, Motilal Nehru National Institute of Technology Allahabad, Prayagraj, India

structure) [2]. The following five types of sentence structures have been investigated in the literature using fMRI data: violations of syntactic rules, sentences with and without pseudo-words, sentences with different agreements and types, sentences versus word lists, and sentences versus simple sentences.

New insights have emerged from the use of fMRI data in the study of syntax, or sentence structure, in the brain [3]. When sentences are grammatically incorrect, known as syntactic violations, areas involved in syntactic processing experience an increase in activation. This is because agreement checking and structure building are disrupted, resulting in more attentive behavior [4]. For syntactic errors, the brain's superior frontal region is more active than for other errors. The brain's activation patterns were different when grammatically correct and incorrect sentences were compared. More brain regions, including Broca's area and various regions of the left temporal lobe, are activated by complex sentences, which require more syntactic operations [5]. The anterior part of the temporal lobe, particularly the temporal pole, as well as the superior and middle temporal gyri, show increased activation when compared to sentences and a list of unrelated words [6]. The posterior and anterior superior temporal sulcus was shown to be activated in the comparison between simple sentences and sentences with words that have no meaning (pseudo-words) [7].

By altering the structure and agreements of various types of sentences, researchers have looked into how our brains process them [8]. According to a study that is referred to as [9], processing the genitive case in the English language results in greater activation of the left inferior frontal gyrus and the back part of the middle temporal gyrus than the processing of the nominative and accusative cases. This study demonstrated how the brain processes various language cases.

Due to the additional syntactic component for negation, studies have demonstrated that negative sentences have a more complex syntactic structure than affirmative sentences [10]. The brain's representation of negative sentences varies depending on whether the sentence has a contradictory or bipolar predicate [11]. While affirmative sentences are thought to be processed directly, negative sentences are thought to be processed in two steps, with the affirmative form being processed first and the negative version being represented in the brain [12]. When processing negative sentences, this additional syntactic transformation causes more brain activity in the left posterior temporal gyrus and bilateral parietal regions than when processing affirmative sentences [13]. Using a target-probe matching task, researchers found that when Japanese–English bilinguals heard negative English sentences, the left temporal and pre-central gyrus were more active [14]. A dynamic model of processing negative sentences has been demonstrated by studies on both action-related and abstract sentence polarity [15], with action-related negative sentences experiencing partial deactivation in the left pallidum while abstract negative sentences experience increased activation in the left perisylvian and parietal regions. Positive sentences were found to be more active in the left premotor cortex (BA6) in another Danish language study [16]. According to the findings of this study, negative sentence processing involves three primary brain regions: the left inferior frontal gyrus is used to calculate syntactic complexity, the bilateral inferior parietal regions are used for semantic processing, and the left premotor cortex is used for sentence structure. A study using functional magnetic resonance imaging (fMRI) revealed that affirmation activates the right supramarginal gyrus, which is associated with semantic processing, while negation activates the left premotor cortex, which is associated with rule-based memory processing. Using functional magnetic resonance imaging (fMRI), a study of German sentences that contained both single and double negations was carried out [17]. The findings demonstrated a connection between the processing of main clause negations and a number of brain regions, including the left pars triangularis, left pars opercularis, left supplementary motor area, and left superior temporal gyrus. The most important part of coordinating the brain regions involved in language processing and logical reasoning was done by the inferior frontal gyrus. Using fMRI, another study looked at how the brain responds to affirmative and negative Hindi sentences [18]. The findings demonstrated that the processing of affirmative and negative sentences was frequently mediated by the bilateral inferior frontal gyrus, left parietal cortex, left fusiform, bilateral supplementary motor area, bilateral temporal gyrus, and bilateral occipital area. Negative sentence processing also occupied the anterior temporal pole. The sentence processing fMRI data from the star-plus dataset was analyzed using a feature selection method [19]. The star-plus dataset is used in different types of analysis, such as the cognitive state of the brain [20], MVPA pattern analysis of brain signals [21], deep knowledge representations [22], and many more. A NBTree classifier was used to select the data, classify the brain state, and determine whether a person was reading an affirmative or negative sentence. The model's accuracy rate was determined to be 92.91 percent by the outcomes. The proposed method was aimed at determining whether the brain was in a positive or negative state while processing sentences. The model had an accuracy rate of 100%, according to the findings by Rotation Forest and SVM-RFE. The model's procedures and components are described in the next section, followed by results and

analysis in the third section, and a conclusion in the last section.

Methods and Model

The task of detecting the polarity of a sentence can be broken down into seven steps. First, fMRI data are collected from participants while they are engaged in a specific task, such as viewing pictures or sentences on a screen. These data are then pre-processed to eliminate any artifacts. The focus of the analysis is to study the brain's behavior during sentence processing, so data related to sentence processing is extracted. Using a correlation-based feature selection method and a decision tree classifier, the model can accurately classify the brain's state during sentence polarity detection. The proposed approach is illustrated in a flow diagram (Fig. 1), with each step explained in detail in subsequent sections.

fMRI Data Collection

A 3.0 T GE Sigma scanner and 500 ms intervals were used to collect MRI data from six healthy subjects for this study. The scan took images of approximately 5000 voxels per subject across eight oblique axial slices divided into two separate four-slice volumes with a TE of 18 ms and a flip angle of 50° [23].

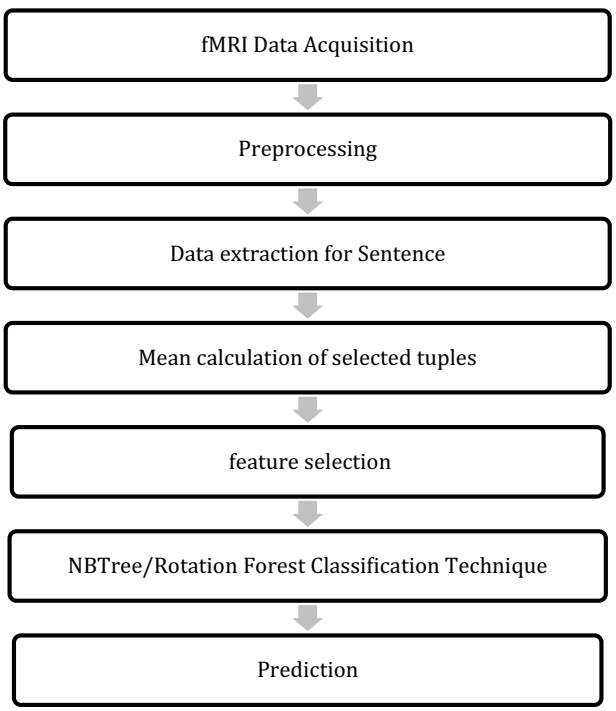

Fig. 1 Working steps of proposed model

Pre-processing

During the fMRI data collection, a 3.0 T GE Sigma scanner with a flip angle of 500 and a TE of 18 ms was used to scan six healthy individuals. Each person's 8 slices contained approximately 5000 voxels, which were divided into two volumes of four slices. The images were then processed with the FIASCO program [24] to align the slice timing and get rid of artifacts such as head motion and signal drift. The Talairach co-ordinate system was used to normalize the data. The calcarine sulcus (CALC), the left and right dorsolateral prefrontal cortex (LDLPFC, RDLPFC), the left and right frontal eye fields (LFEF, RFEF), the left and right intraparietal sulcus (LIPS, RIPS), the left and right inferior frontal gyrus (LIFG, RIFG), supplementary motor areas (SMA), and the left and right opercularis (LOPER, ROPER).

Data Extraction for Sentence Processing

The fMRI signals associated with sentences are extracted from the dataset in order to investigate the brain's response to sentence processing. There are two datasets in the StarPlus collection: the PS dataset and the SP dataset. The sentence first appeared on the screen at $t=8$ s and lasted for 4 s in the PS dataset. The picture was shown first. The sentence was the first stimulus in the SP dataset. It appeared on the screen at $t=0$, lasted for four seconds, and then there was a rest period of four seconds. From $t=0$ to $t=8$, 16 brain images were extracted for the SP dataset. In a similar vein, 16 images were extracted from the PS dataset from $t=8$ to $t=16$. Since fMRI BOLD signals typically last 9–12 s beyond the neural activity of interest, despite the fact that the sentence stimulus was only presented for 4 s, an 8 s interval of fMRI data was used to train the classifier. The feature selection entity then uses these images to determine the most prominent stimuli features.

Mean Calculation

The extracted data for classifier training take the form

$$f : fMRI - \text{Sequence}\,(t, t+8) \rightarrow \{\text{Affirmative/Negative sentence}\} \quad (1)$$

where $t=0$ for SP dataset and $t=8$ for PS dataset.

The dataset obtained after extracting the fMRI signals corresponding to the sentences is very large to analyze. For each subject, the resulting matrix for one trial of SP or PS data is 16 rows by 5000 columns. To reduce the size of data, the mean of all the 16 images for each feature vector point is calculated. This results in a matrix with 1

row by 5000 columns for one trial, instead of 16 rows by 5000 columns. By repeating this process for all 40 trials (20 positive and 20 negative sentences), an array with 40 rows by 5000 columns is obtained for a single subject. The mean is calculated across all column entries using a specific formula:

$$\forall j\, j \in \text{column}, \sum_{i=1}^{n} x_{ij}/n \text{ Where } x_{ij} \text{ is an } ith \text{ row and } jth \text{ column entry} \quad (2)$$

and $1 \leq j \leq number.column(\approx 5000)$

Feature Selection

CFS (Correlation-Based Feature Selector) Filter Algorithm

There are three steps involved in selecting features: attribute selection mode, search method, and attribute evaluator. By removing irrelevant and redundant features, the objective is to identify a subset of features that are both minimal and collectively predictive. In a correlation-based feature selector, the attribute evaluator chooses a subset of highly correlated but low intra-correlation features. The CFS (correlation-based feature selector) filter algorithm uses the correlation heuristic evaluation function to rank feature subsets. Since they are less correlated with the class, irrelevant features are eliminated, whereas redundant features are eliminated because they are highly correlated with one or more features. The definition of the CFS feature subset evaluation function [25] is

$$M_s = \frac{kr_{cf}}{\sqrt{k + k(k-1)r_{ff}}} \quad (3)$$

where M_s is the heuristic "merit "of a feature subset S containing k features, r_{cf} is the mean feature class correlation $f \in S$ and r_{ff} is the average feature–feature inter-correlation.

The equation measures how well a set of features predicts the class (numerator) and the level of redundancy among those features (denominator). The search method uses a greedy approach [26], moving forward and backward through the possible feature subsets to find the best combination. The feature that increases the predictive performance the most during the forward iteration is selected. This continues until no more improvement is possible or a set limit is reached. In the backward iteration, features that do not harm performance are removed until no more can be taken out without affecting performance. The correlation-based subset evaluator feature selection algorithm uses the selected features (SF), target data (TD), and a set of variables (VD) to determine the optimal feature subset.

Algorithm-(Forward–Backward Selection)

Input: InputData I, TargetData TD, Significance Level L

1. Set Selected Features (SF) to an empty set
2. Begin the forward phase:
a. While SF changes:
i. Identify the variable with the minimum p-value (V*) in relation to SF and Target Data (TD)
ii. If the p-value of TD and V* given SF is less than or equal to a threshold (L), then add V* to SF
3. Begin the backward phase:
a. While SF changes:
i. Identify the variable with the maximum p-value (V*) in relation to SF and Target Data (TD)
ii. If the p-value of TD and V* given SF is greater than the threshold (L), then remove V* from SF
4. Return SF as the final set of selected features

Attributes are selected over the entire dataset to make the selection justifiable. After the attribute selection phase is over, classification is done.

SVM-RFE (Support Vector Machine–Recursive Feature Elimination)

SVM-RFE (Support Vector Machine–Recursive Feature Elimination) [29] is a feature selection technique in which

a support vector machine (SVM) model is used to identify the importance of each feature in a dataset. In this approach, the least important features are recursively removed until the desired number of features is reached or the performance of the model is optimized. This method is particularly useful in reducing the dimensionality of high-dimensional data, removing noise, and improving the performance of machine-learning models. The SVM-RFE approach is often used in the fields of computer vision, natural language processing, and bioinformatics, among others. SVM Recursive Feature Elimination (RFE) is a feature selection algorithm that uses Support Vector Machine (SVM) as the base classifier. The aim of this algorithm is to rank the features based on their contribution to the classification task and select the most important ones.

Here is a general outline of the SVM-RFE algorithm:

1. **Initialize the number of features to be selected, "k"**
2. **Train an SVM classifier on the entire set of features**
3. **Compute the feature weight (coefficient magnitude) for each feature**
4. **Sort the features based on their weight**
5. **Repeat the following steps until "k" features are left:**
   a. **Remove the feature with the lowest weight**
   b. **Train an SVM classifier on the remaining features**
   c. **Compute the feature weight for each feature**
   d. **Sort the features based on their weight**
6. **Return the set of "k" features with the highest weight**

The SVM-RFE algorithm uses the magnitude of the feature weights to determine the importance of each feature. The idea is that the most important features will have the highest weights, and the least important features will have the lowest weights. The SVM-RFE algorithm has been found to be effective in reducing the feature space while maintaining the performance of the model. This feature selection approach is particularly useful when dealing with high-dimensional datasets where the number of features is larger than the number of samples.

Classification

NBTree Algorithm

The utility of each attribute in the dataset is the first step in the NBTree algorithm, which is a hybrid of decision tree and Naive Bayes. The attribute with the greatest utility is chosen and evaluated to see if it outperforms the current node significantly. The algorithm ends if it is not, and a Naive Bayes Classifier is created for the current node. If it is superior, the dataset is partitioned based on the value of each attribute, and the algorithm is called recursively on each partition until a Naive Bayes Classifier is created for a node or the utility of the attributes is no longer significant.

A dataset has been analyzed and categorized using the NBTree algorithm [27]. It is a decision tree and Naive Bayes in one. At each node, the algorithm conducts a Naive Bayes analysis on each of the five equal parts of the data. The Naive Bayes classifier's error rate is used to calculate the split. This algorithm consists of two steps: the stage of learning and the stage of classification. During the learning phase, a model with the appropriate class labels is constructed from the training data. This built model is used to classify new, unobserved data during the classification stage. A set of labeled instances is the algorithm's input, and the decision tree with Naive Bayes categorization at the leaves is the algorithm's output.

Algorithm:
1. Calculate u(Xi) to determine how useful it is to split each attribute Xi. Set a threshold for continuous attributes as well.
2. j = arg maxi, choose the attribute with the greatest utility.
3. Create a Nave-Bayes classifier for the current node and return if the selected attribute is not significantly more useful than the current node.
4. Divide the data set T in accordance with the Xj test. A threshold split is used for continuous attributes, while a multi-way split based on all possible values is used for discrete attributes.
5. Call the algorithm recursively on the portion of T that matches the test leading to that child for each child produced by the split.

Rotation Forest Classifier

The Rotation Forest classifier is an ensemble machine-learning algorithm that combines the strengths of decision trees and principal component analysis (PCA). In a Rotation Forest, each individual tree in the forest is grown from a rotated version of the original feature set, created using PCA. This rotation allows each tree to focus on different aspects of the data and helps to reduce the correlation between trees, leading to improved performance compared to traditional decision trees.

During prediction, each tree in the forest makes a prediction and the final class label is determined based on a majority vote of the trees. This approach helps to reduce overfitting, as each tree is trained on a different version of the feature set, and improves the overall accuracy of the classifier by combining the predictions of multiple trees. The Rotation Forest classifier has been shown to outperform other machine-learning algorithms in several domains, including image classification and text classification.

The algorithm of the Rotation Forest classifier [28] can be summarized as follows:

1. Initialize the number of trees in the forest, "ntrees"
2. For each tree in the forest:
a. Randomly select "d" features from the total set of features
b. Compute the principal component analysis (PCA) of the selected features
c. Rotate the PCA to obtain the new feature space
d. Split the data into "k" clusters using the k-means algorithm
e. For each cluster:
i. Select a random subset of the data points
ii. Build a decision tree using the subset
f. Average the prediction of all decision trees in the cluster to obtain the final prediction
3. Return the average prediction of all trees in the forest as the final prediction
The parameters "d" and "k" can play a crucial role in determining the performance of the Rotation Forest algorithm.

Table 1 Number of voxels in each anatomical region of the brain

ROI	Subject 04799	Subject 04820	Subject 04847	Subject 05675	Subject 05680	Subject 05710
LPPREC	13	190	153	26	53	110
RPPREC	59	124	144	50	72	43
RSGA	71	41	34	54	71	57
RFEF	81	142	68	45	40	73
ROPER	88	140	180	194	255	108
LFEF	97	128	109	30	71	124
LSGA	104	55	7	119	95	127
RIPS	118	65	166	158	237	185
RTRIA	144	213	57	146	42	118
LOPER	145	103	169	261	235	101
LIPS	147	155	235	197	196	226
SMA	154	229	215	77	71	155
LSPL	171	290	308	265	274	137
LTRIA	190	175	113	136	93	150
RSPL	213	238	252	254	213	119
CALC	255	408	318	399	334	219
RIT	285	187	277	267	151	209
RIPL	287	46	90	244	215	213
LIPL	299	62	133	263	286	264
LT	340	484	305	476	472	445
RT	356	440	286	325	346	357
LIT	366	225	286	233	173	197
RDLPFC	468	374	349	382	453	419
LDLPFC	498	501	440	504	614	478
Whole brain voxels	4949	5015	4698	5135	5062	4634

The data for sentence processing are taken from the Star-Plus dataset using Eq. 1

Table 2 Subject-wise selected voxels

Subject id	Selected attributes
04799	59, 129, 188, 198, 316, 326, 370, 392, 426, 434, 449, 546, 553, 732, 855, 866, 944, 971, 1042, 1068, 1090, 1366, 1446, 1452, 1470, 1532, 1551, 1566, 1742, 1803, 2274, 2401, 2402, 2501, 2505, 2803, 2876, 2889, 2896, 2923, 2947, 2956, 2993, 3113, 3159, 3213, 3271, 3286, 3341, 3390, 3403, 3417, 3594, 3609, 3614, 3630, 3657, 3706, 3908, 4105, 4390, 4432, 4717, 4758: 64
04820	291, 331, 348, 1156, 1361, 1369, 1476, 1542, 1571, 1609, 1692, 1722, 1750, 1873, 1888, 1925, 1932, 1933, 1934, 2000, 2075, 2239, 2422, 2552, 2892, 2909, 3140, 3182, 3278, 3397, 3414, 3430, 3544, 3576, 3614, 3783, 3810, 3846, 3877, 4082, 4162, 4173, 4311, 4349, 4398, 4435, 4459, 4698, 4841: 49
04847	29, 67, 148, 202, 281, 342, 441, 520, 573, 899, 917, 1185, 1205, 1212, 1307, 1440, 1453, 1493, 1566, 1604, 1616, 1660, 1740, 1917, 2034, 2088, 2094, 2226, 2548, 2891, 2901, 2984, 3114, 3340, 3535, 3545, 3645, 3713, 3716, 3773, 3881, 4030, 4180, 4319, 4333, 4380, 4385, 4412, 4547, 4582, 4684: 51
05675	208, 221, 364, 438, 441, 456, 524, 758, 817, 859, 932, 950, 991, 1007, 1065, 1241, 1282, 1535, 1561, 1565, 1583, 1617, 1653, 1656, 1790, 1824, 1849, 1875, 1992, 2059, 2164, 2252, 2306, 2384, 2393, 2453, 2573, 2601, 2633, 2975, 3021, 3059, 3097, 3102, 3123, 3137, 3138, 3197, 3532, 3580, 3621, 3827, 3840, 3932, 4086, 4390, 4538, 4567, 4585, 4696, 4845: 61
05680	45, 123, 611, 786, 935, 939, 1051, 1217, 1342, 1351, 1378, 1449, 1492, 1541, 1699, 1866, 1915, 1990, 2021, 2098, 2135, 2281, 2337, 2487, 2516, 2531, 2539, 2557, 2883, 2891, 2912, 2959, 3229, 3285, 3331, 3362, 3690, 3697, 3756, 3879, 3924, 4008, 4075, 4081, 4161, 4176, 4448, 4592, 4772, 4920, 5048, 5061: 52
05710	35, 242, 344, 529, 598, 711, 751, 832, 847, 886, 1075, 1252, 1277, 1479, 1488, 1542, 1552, 1560, 1756, 1938, 2091, 2233, 2379, 2441, 2643, 2792, 2967, 3009, 3038, 3075, 3431, 3542, 3622, 3656, 3693, 3715, 3807, 3841, 4070, 4118, 4165, 4253, 4261, 4309: 44

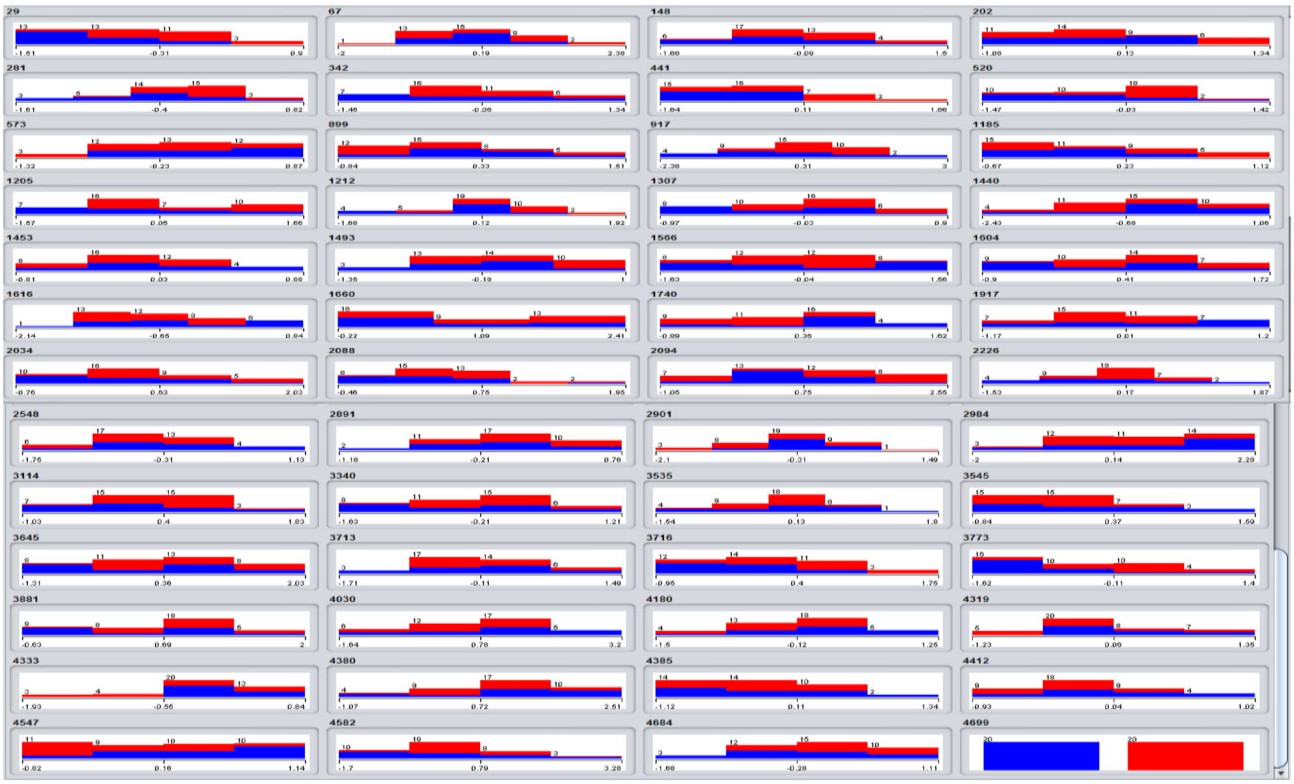

Fig. 2 Statistical distribution and graphical representations of selected features

Table 3 Subject-wise selected attributes (voxel)

Area	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6
LOPER	0	1	3	0	5	0
LPPREC	0	2	0	1	1	3
RSGA	0	1	2	0	1	0
RTRIA	0	6	1	0	0	0
LSGA	1	1	2	2	0	1
ROPER	1	1	2	2	3	0
RPPREC	1	0	0	1	0	0
LIPS	2	0	2	3	0	1
LIT	2	1	3	5	3	5
LSPL	2	3	2	1	3	1
LTRIA	2	1	1	0	1	2
RFEF	2	2	0	1	0	1
RSPL	2	2	6	4	2	2
SMA	2	3	2	1	0	1
LDLPFC	3	5	3	5	6	5
LFEF	3	3	1	0	1	2
LIPL	3	1	2	10	7	0
LT	3	3	4	3	4	4
RIPS	3	0	0	3	1	2
RIT	4	1	1	2	0	2
RT	4	5	7	2	3	3
RDLPFC	6	3	3	3	6	4
CALC	9	4	4	8	3	2
RIPL	9	0	1	4	2	3

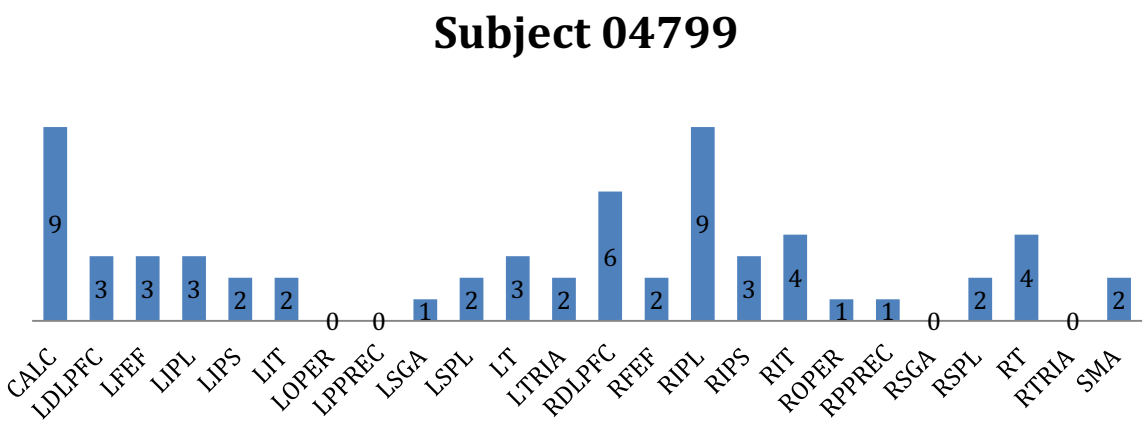

Fig. 3 Region-wise number of selected feature vectors for subject 04799

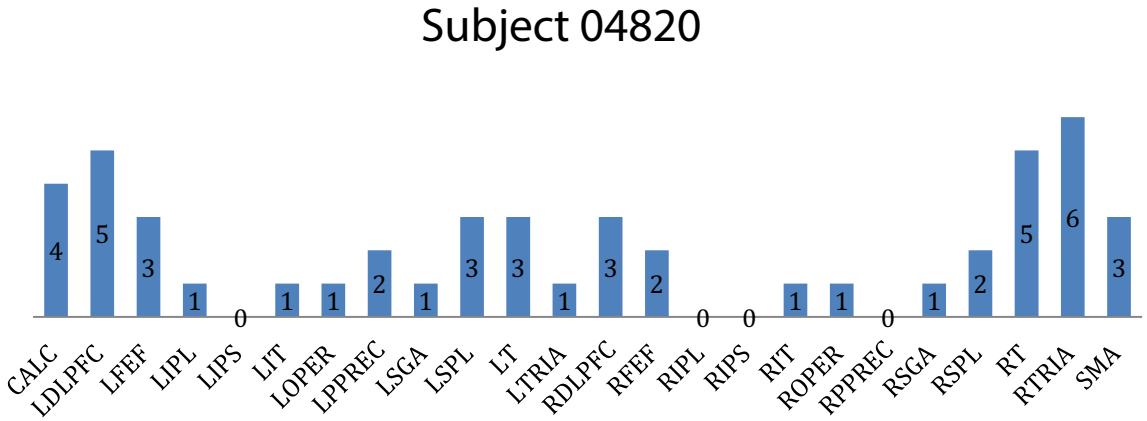

Fig. 4 Region-wise number of selected feature vectors for subject 04820

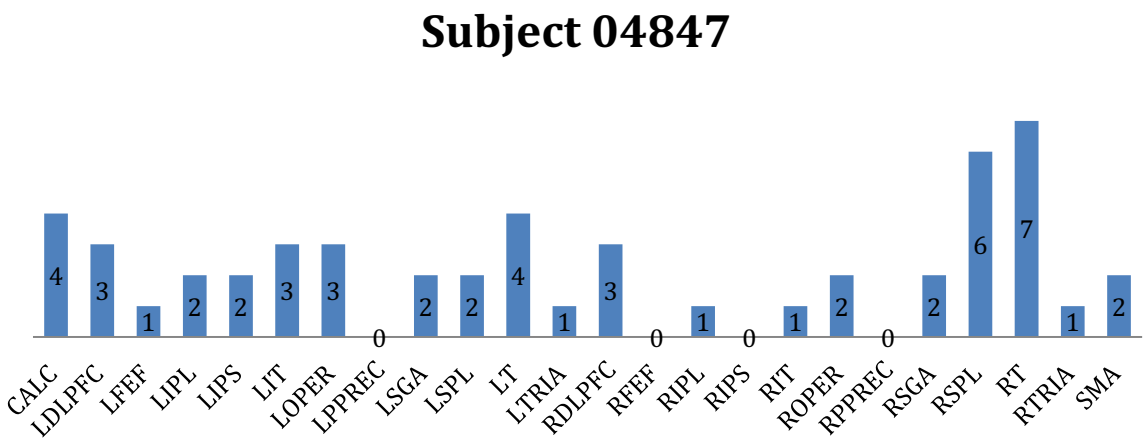

Fig. 5 Region-wise number of selected feature vectors for subject 04847

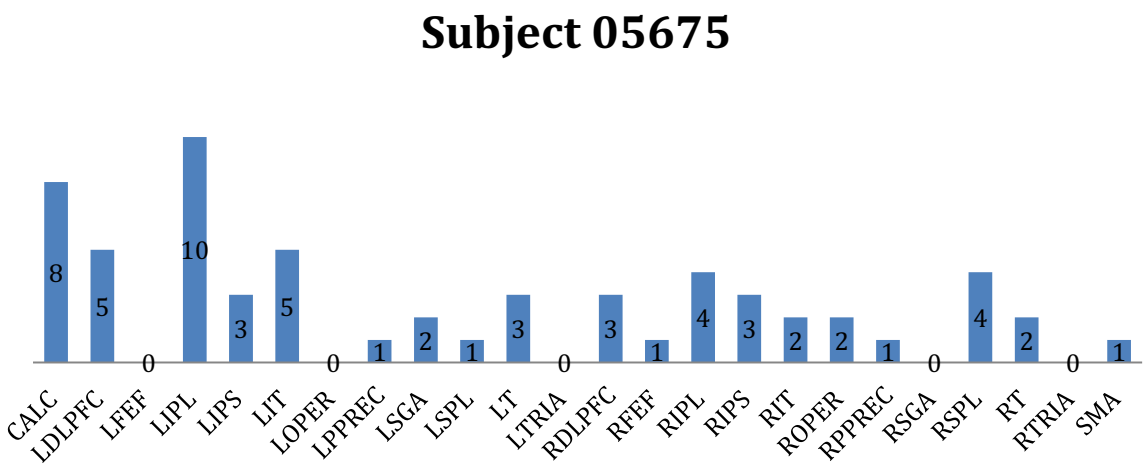

Fig. 6 Region-wise number of selected feature vectors for subject 05675

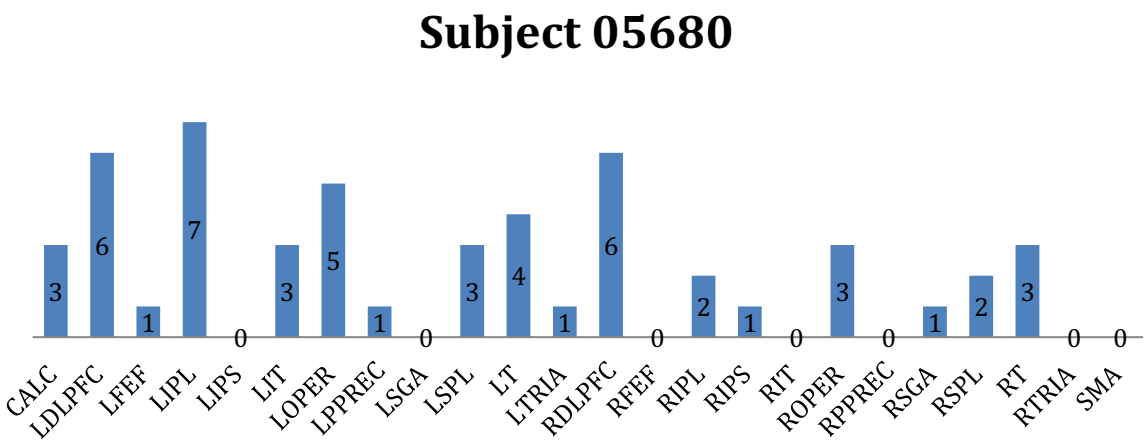

Fig. 7 Region-wise number of selected feature vectors for subject 05680

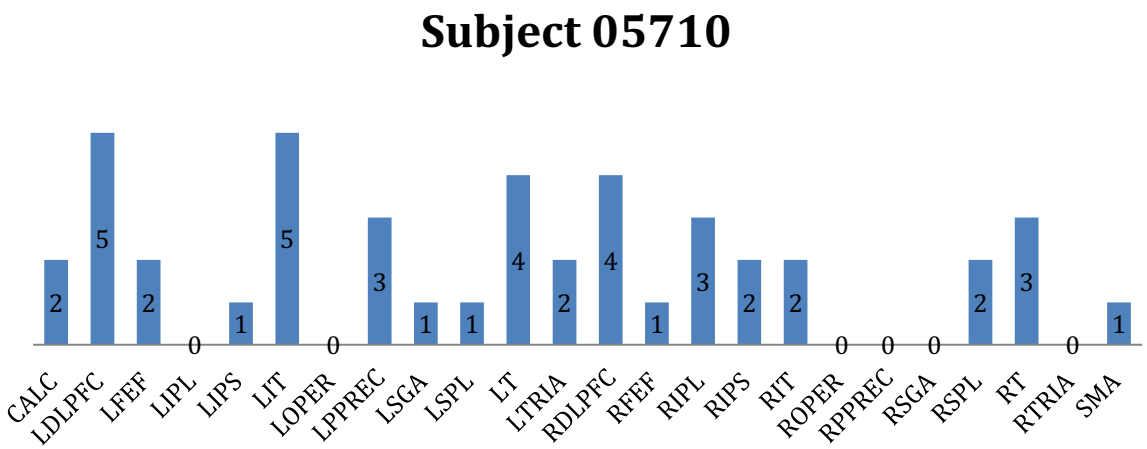

Fig. 8 Region-wise number of selected feature vectors for subject 05710

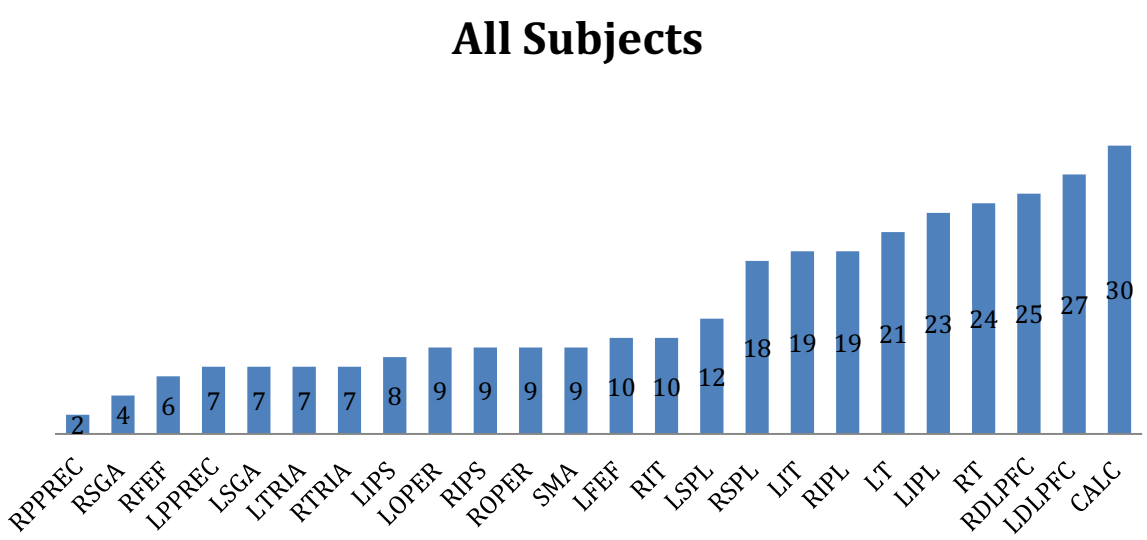

Fig. 9 Total number of features selected region wise from all subjects

Table 4 Result analysis table for all six subjects

Subject Id	TP rate	FP Rate	Precision	Recall	F-measure	ROC area	Accuracy
04799	0.975	0.025	0.976	0.975	0.975	0.988	0.975
04820	0.975	0.025	0.976	0.975	0.975	0.988	0.975
04847	0.825	0.175	0.832	0.825	0.824	0.823	0.825
05675	0.975	0.025	0.976	0.975	0.975	0.957	0.975
05680	0.825	0.175	0.87	0.825	0.819	0.988	0.825
05710	1.000	0.000	1.000	1.000	1.000	1.000	1.000

Result Analysis and Discussion

Dataset Description

Marcel Just and his group at Carnegie Mellon University gathered the Star-Plus dataset for an fMRI study. The fMRI machine recorded the brain activity of six healthy subjects as they performed a series of tasks in the experiment. The experiment lasted 27 s for each trial, and brain activity was recorded every 500 ms. In a single trial, 54 images of the brain were recorded.

On a screen, subjects were shown pictures and sentences in the experiment. After that, a button was pressed to let them know if the sentence correctly described the picture. During the trial, the presentation order was shifted to either picture first, followed by the sentence, or picture first, followed by a sentence. There were two categories of trials: PS stands for the picture that was shown first, and SP stands for the sentence that was shown first.

In the experiment, pictures of geometrical arrangements of symbols such as $, *, and + were shown. Half of the sentences were positive and half were negative, and they were descriptions of the picture. 40 trials with approximately 5000 voxels each were presented to each subject, totaling 270,000 voxels in a single trial. In order to discover distinct

Table 5 Confusion matrix of all subjects

Subject id	Predicted class		Actual class
	Affirmative	Negative	
04799	19	1	Affirmative
	0	20	Negative
04820	19	1	Affirmative
	0	20	Negative
04847	15	5	Affirmative
	2	18	Negative
05675	19	1	Affirmative
	0	20	Negative
05680	20	0	Affirmative
	7	13	Negative
05710	20	0	Affirmative
	0	20	Negative

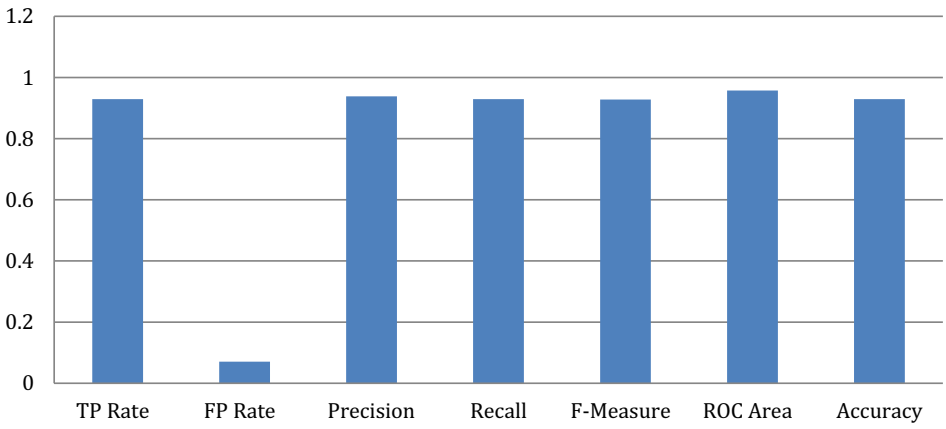

Fig. 10 Average of performance analysis of polarity detection

patterns of voxel activity, the brain was divided into 25 distinct regions of interest (ROIs).

Results and Discussion

The study focuses on analyzing the distribution of active brain voxels using statistical methods. The important voxels are selected through feature selection and then classified using NBTree. The results are then compared with other existing methods on the same data. The data were collected from six individuals, each with a unique identifier: 04799, 04820, 04847, 05675, 05680, and 05710. The dataset includes fMRI signals from 24 different regions of the brain. The analysis of the dataset shows that the number of active voxels varies not only by region but also by the individual. Table 1 is presented showing the distribution of signal data in different regions of the brain.

The average of these data is found using Eq. 2 and is used to study the brain activity during the sentence polarity task. To eliminate unnecessary features and find the most important ones, we applied the correlation-based subset evaluator feature selection method. The resulting set of features is highly correlated and the subject-specific attribute values are listed in Table 2.

Figure 2 displays the statistical distribution of chosen feature values for subject 04847. The illustration demonstrates that the selected features have a spread-out distribution and are not biased to one side. The distribution does not lean to the right or left and most of the selected features are unique and distinguishable from each other. This even distribution of less redundant selected features confirms the accuracy of our feature selection process as it has only picked important features. The blue color denotes the distribution of affirmative sentence and the red color denotes the distribution of negative sentences and their corresponding numeric values represents the statistical distribution of a feature for affirmative or negative samples. For a particular voxel, there are 40 data points of which 20 are for affirmative sentences and the rest 20 are for negative sentences. The graphical view represents the activity pattern of different data points of a particular voxel in both types of cases.

Since it varies from person to person, the frequency of particular features in various brain regions does not provide a clear indication of which brain region plays a significant role in the sentence polarity classification task. It was discovered that some regions have a greater impact on determining the polarity of sentences, while others have little to no impact, by calculating the sum of active voxels in each region from all subjects. CALC, LDLPF, and RDLPFC were found to have the greatest impact on categorizing affirmative and negative sentences, while RPPREC, RSGA, and RFEF were found to have the least impact. The subject-wise features chosen for the sentence polarity classification task across all brain regions are shown in Table 3.

Table 6 Result obtained without feature selection from all subjects

Subject id	TP rate	FP rate	Precision	Recall	F-measure	ROC area	Accuracy
04799	0.55	0.45	0.551	0.55	0.549	0.587	0.55
04820	0.45	0.55	0.449	0.45	0.449	0.491	0.45
04847	0.40	0.60	0.399	0.40	0.398	0.27	0.40
05675	0.575	0.425	0.575	0.575	0.575	0.589	0.575
05680	0.450	0.550	0.445	0.450	0.437	0.497	0.45
05710	0.675	0.325	0.675	0.675	0.675	0.697	0.675

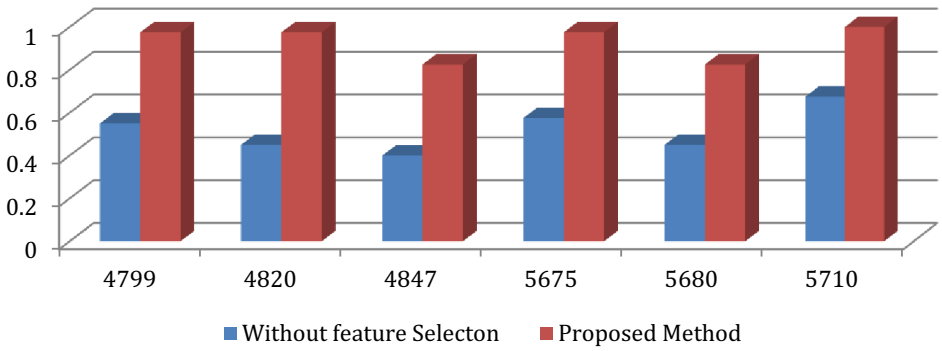

Fig. 11 Subject-wise comparative analysis of polarity detection accuracy

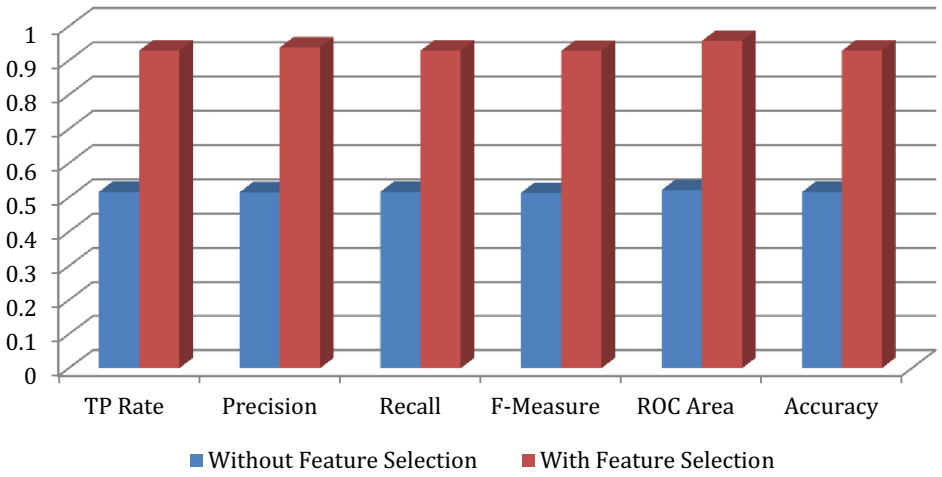

Fig. 12 Comparative analysis (overall average of six subjects)

Table 7 The result of SVM-RFE + rotation forest

Subject Id	TP rate	FP rate	Precision	Recall	F-measure	ROC area	Accuracy
04799	1.000	0.000	1.000	1.000	1.000	1.000	1.000
04820	1.000	0.000	1.000	1.000	1.000	1.000	1.000
04847	1.000	0.000	1.000	1.000	1.000	1.000	1.000
05675	1.000	0.000	1.000	1.000	1.000	1.000	1.000
05680	1.000	0.000	1.000	1.000	1.000	1.000	1.000
05710	1.000	0.000	1.000	1.000	1.000	1.000	1.000

Table 8 Feature vs. accuracy

No. of features	10	20	50	100	200	400	1000	2000	5000
Accuracy	100	100	100	98.6	97.5	93.2	85.3	73.2	52.4

The plot below displays the frequency of voxels selected as features for each subject. The number of selected feature voxels from various brain regions is plotted in Figs. 3, 4, 5,6, 7, and 8 for subjects 1 to 6, respectively.

The total number of voxels selected for each brain region in determining sentence classification for all 6 subjects is: CALC (30), LDLPFC (27), LFEF (10), LIPL (23), LIPS (8), LIT (19), LOPER (9), LPPREC (7), LSGA (7), LSPL (12), LT (21), LTRIA (7), RDLPFC (25), RFEF (6), RIPL (19), RIPS (9), RIT (10), ROPER (9), RPPREC (2), RSGA (4), RSPL (18), RT (24), RTRIA (7), and SMA (9). The data are shown in Fig. 9 and it shows that the most significant

regions for sentence polarity classification tasks are CALC, LDLPFC, and RDLPFC, while RPPREC, RSGA, and RFEF have minimal impact. The other brain regions contribute very little to the selection of prominent feature voxels. The result is evident as the largest number of voxels from the entire brain data is found in CALC, LDLPFC, and RDLPFC.

A set of selected features were used by the authors to train a classifier. The classifier was taught to distinguish between positive and negative sentence processing patterns of brain activity. They assumed that the machine-learning algorithm could learn the temporal pattern that differentiates between positive and negative sentences and that the fMRI data contained sufficient information to reveal the brain state associated with a particular sentence type. The selected features were categorized using the NBTree method. Whether positive or negative sentences were being processed, the model was able to accurately identify the state of the brain. Table 4 provides a summary of the findings, and Table 5 contains the confusion matrix for each subject's findings. With an accuracy of over 92 percent, the model was trained and tested using tenfold cross-validation. The average overall performance analysis is depicted in Fig. 10, and the subject-wise accuracy results for the sentence polarity task are presented in Table 4. The figure displays the performance metrics, including accuracy, which indicate an approximate 93% classification accuracy on average.

- The performance metrics used in the analysis of the results have standard definitions:
- TP Rate refers to the rate of correctly classified instances as a particular class.
- FP Rate represents the rate of instances that have been wrongly classified as a specific class.
- Precision refers to the proportion of correct instances classified as a certain class over the total number of instances classified as that class.
- Recall represents the proportion of instances classified as a specific class divided by the total number of actual instances in that class. It is similar to TP Rate.
- F-Measure is a measure of the balance between precision and recall, calculated as 2*Precision*Recall/(Precision + Recall).
- The ROC (receiver-operating characteristics) area shows the overall performance of the classifier.

The success of a classifier is determined by the number of incorrect predictions made by it, also known as the error rate. The purpose of the classifier is to accurately predict the class of instances it is presented with. The confusion matrix provides insight into the accuracy of the model by showing the number of true positive and false positive predictions, among others. The table displays the correct and incorrect predictions made by the classifier by comparing the predicted class values shown in the columns with the actual class values shown in the rows.

Table 9 Efficiency of classifier

Methods	Kappa statistic	Root means squared error	Root relative squared error
Random Forest	0.9	0.2208	44.1584%
k-NN	0.75	0.3451	69.0257%
MLP	0.85	0.2259	45.1868%
Rotation Forest	0.95	0.2002	40.4612%

Table 10 Comparing feature selection methods

Result analysis	1	2	3	Total
N	6	6	6	18
$\sum X$	5.725	2.915	5.575	14.215
Mean	0.9542	0.4858	0.9292	0.79
$\sum X^2$	5.4644	1.4354	5.1869	12.0867
Std.Dev	0.0188	0.062	0.0368	0.225

Comparative Analysis

The comparison of the results between Tables 4 and 6 shows that using the feature selection technique, the accuracy of the classification has improved by 40%. This dispels the notion that applying feature selection leads to a loss of information and a decrease in results. In fact, our results indicate the opposite. The comparison between the proposed approach and the approach without feature selection can be seen in Fig. 11, which displays the subject-wise results.

The results obtained from using the proposed method, which includes feature selection, were compared with the results without feature selection in Fig. 12. It was found that the overall average results for the proposed method were significantly better than the method without feature selection. This was reflected in the improved TP rate, precision, recall, F-Measure, ROC area, and accuracy.

Rotation Forest + SVM-RFE

With SVM-RFE feature selection and Rotation Forest Classification, we obtained 100% accuracy in classification. Although this accuracy shows that there is an overfitting problem. Many binary class classification problems have been solved with 100% accuracy with new machine-learning development. The results obtained are summarized in Table 7. We have tested and retested our results with different perspectives to justify the results.

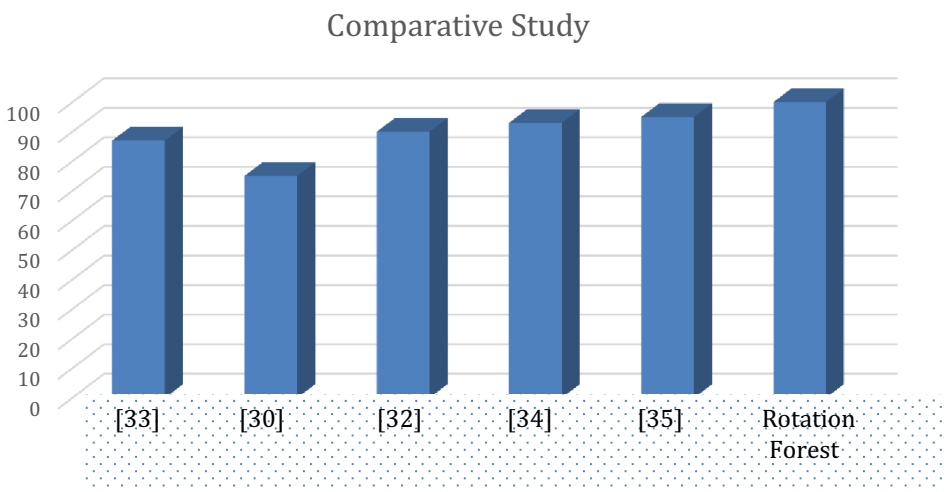

Fig. 13 Comparative accuracies of recent works on the same dataset

Initially, we tested our results on default parameters only and that gave 100% accuracy. We tested the data again with SVM-RFE and Rotation Forest with the detailed analysis, and we achieved the following results:

1. Based on no. of features fed to the classifier in Table 8.
2. Based on test method with k-fold cross-validation with 100 features, we got 98.6% accuracy. But with a split of 70% data for training and 30% for testing, we got 97.2% accuracy.
3. However, when we tested the combination of these methods on other data such as Noun Classification and Reading Harry Potter Chapter, we could not achieve more than 70% accuracy. This shows the SVM-RFE and Rotation Forest are giving significant results with the Star-Plus dataset only.

The reason why Rotation Forest may give better accuracy than Naive Bayes Tree can be attributed to several factors. First, Rotation Forest has the advantage of being an ensemble method, which tends to have better generalization performance than individual models. The random feature subsets used in Rotation Forest help to increase the diversity of the model and reduce the risk of overfitting. Second, Naive Bayes Tree assumes that the features are independent, which may not be true in many real-world datasets. This assumption can lead to inaccurate predictions when there are strong correlations between the features. In contrast, Rotation Forest does not make any assumptions about the relationships between the features and can capture more complex interactions. Finally, the performance of any machine-learning algorithm depends on the characteristics of the dataset and the pre-processing steps used. In some cases, Naive Bayes Tree may perform better than Rotation Forest, and vice versa. It is important to experiment with different algorithms and evaluate their performance on a variety of datasets to find the best model for a specific problem.

Rotation Forest and Random Forest are both ensemble learning methods used for classification and regression. The difference between the two lies in the way the trees are generated. In Random Forest, multiple decision trees are generated by randomly selecting a subset of features for each tree, whereas in Rotation Forest, each decision tree is generated by first rotating the feature space using PCA (Principal Component Analysis) or ICA (Independent Component Analysis) and then selecting a subset of features for each tree.

The reason why Rotation Forest can give better results than Random Forest is because the rotated feature space in Rotation Forest helps to eliminate correlations between features and reduce the curse of dimensionality. This results in improved generalization performance, as well as reduced overfitting. In addition, Rotation Forest also helps to improve the accuracy of predictions by combining the results of multiple rotated trees, rather than just combining the results of multiple trees with randomly selected features.

However, it is important to note that the relative performance of Rotation Forest and Random Forest will depend on the specific dataset and problem being solved. In some cases, Random Forest may perform better, while in others, Rotation Forest may be the preferred choice.

Analysis of Results

We have calculated, and compared Kappa statistics and root means squared error in the determination of the efficiency of a particular classification method depicted in Table 9.

Comparing the above statistics, the Random Forest was found to be a more efficient classification method.

To the determination of the efficiency of the feature selection method, we performed an ANOVA test using Tukey's HSD (honestly significant difference) among different feature selection results, and we obtained the following results as in Table 10.

Comparison with Results Obtained on the Same Dataset

In comparison to previous studies on the same dataset, our results had much higher accuracy in polarity detection. In Fig. 13, the classification accuracy from two recent papers was plotted for comparison. The first paper, by Behroozi et al. [30] achieved an accuracy of 75% using F-score-based feature selection and SVM classification techniques. The second paper [32] used NeuCube [31] to study the neural activity of affirmative and negative sentences and was able to recognize sentence polarity with an accuracy of 90%. On the same data for sentence polarity detection, [33] uses k-NN with 90% accuracy, [34] uses MLP to produce 93.33% accuracy and [35] uses Random Forest and resulting in an accuracy of 95.41%. However, our method achieved an accuracy of 100%. Our results clearly show that our approach has a much higher classification accuracy than previous methods.

Conclusion

This research paper focuses on analyzing fMRI (functional magnetic resonance imaging) data to detect the state of the brain when processing affirmative and negative sentences. The authors have developed a machine-learning model using a feature selection technique based on correlation and NBTree classification algorithm. The result of the model shows that it can accurately classify the cognitive state in sentence polarity detection tasks with an average accuracy of 92.91%. Further, the accuracy of the result is enhanced up to 100% using SVM-RFE and Rotation Forest. The statistical analysis indicates that the contribution of selected feature voxels varies not only between regions of the brain but also between individuals. The results of the experiment highlight the important role of CALC (cingulate gyrus), RDLPFC (right dorsolateral prefrontal cortex), and LDLPFC (left dorsolateral prefrontal cortex) regions in detecting sentence polarity.

Data Availability Thedataset analysed during the current study are available in the public repository, and the link for the same is: http://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-81/www/.

Declarations

Conflict of interest The authors declare they have no conflict of interest.

References

1. Pandey P, Jha BK, Sinha N. Analyzing cognitive states using fMRI data. Procedia Comp Sci. 2016;90:35–41.
2. Sair HI, Agarwal S, Pillai JJ. Application of resting state functional MR imaging to presurgical mapping: language mapping. Neuroimag Clin. 2017;27(4):635–44.
3. Kaan E, Swaab TY. The brain circuitry of syntactic comprehension. Trends Cogn Sci. 2002;6(8):350–6. https://doi.org/10.1016/s1364-6613(02)01947-2.
4. Fiveash A, Thompson WF, Badcock NA, McArthur G. Syntactic processing in music and language: effects of interrupting auditory streams with alternating timbres. Int J Psychophysiol. 2018;129:31–40.
5. Yang Y, Wang J, Bailer C, Cherkassky V, Just MA. Commonality of neural representations of sentences across languages: predicting brain activation during Portuguese sentence comprehension using an English-based model of brain function. Neuroimage. 2017;146:658–66.
6. Feng S, Qi R, Yang J, Yu A, Yang Y. Neural correlates for nouns and verbs in phrases during syntactic and semantic processing: an fMRI study. J Neurolin. 2020;53: 100860.
7. Meyer L, Friederici AD. Neural systems underlying the processing of complex sentences. In: Neurobiology of language. Academic Press; 2016. p. 597–606.
8. Rogalsky C. The role of the anterior temporal lobe in sentence processing. In: Neurobiology of Language. Academic Press; 2016. p. 587–95.
9. Yokoyama S, Maki H, Hashimoto Y, Toma M, Kawashima R. Mechanism of case processing in the brain: an fMRI study. PLoS ONE. 2012;7(7):e40474. https://doi.org/10.1371/journal.pone.0040474.
10. Haegeman L. The syntax of negation. Cambridge: Cambridge Univerisity Press; 1995.
11. Mayo R, Schul Y, Burnstein E. "I am not guilty" vs "I am innocent": Successful negation may depend on the schema used for its encoding. J Exp Soc Psychol. 2004;40(4):433–49.
12. Zwaan, R. A. (2012). The experiential view of language comprehension: How is negation represented. Higher level language processes in the brain: Inference and comprehension processes, 255
13. Carpenter PA, Just MA, Keller TA, Eddy WF, Thulborn KR. Time course of fMRI-activation in language and spatial networks during sentence comprehension. Neuroimage. 1999;10(2):216–24.
14. Hasegawa M, Carpenter PA, Just MA. An fMRI study of bilingual sentence comprehension and workload. Neuroimage. 2002;15(3):647–60.
15. Tettamanti M, Manenti R, Della Rosa PA, Falini A, Perani D, Cappa SF, Moro A. Negation in the brain: modulating action representations. Neuroimage. 2008;43(2):358–67.
16. Christensen KR. Negative and affirmative sentences increase activation in different areas in the brain. J Neurol. 2009;22(1):1–17.
17. Bahlmann J, Mueller JL, Makuuchi M, Friederici AD. Perisylvian functional connectivity during processing of sentential negation. Front Psychol. 2011;2:104.
18. Kumar U, Padakannaya P, Mishra RK, Khetrapal CL. Distinctive neural signatures for negative sentences in Hindi: an fMRI study. Brain Imaging Behav. 2013;7(2):91–101.

19. http://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-81/www/. Accessed 31 Dec 2022
20. Gupta KO, Chatur PN. Gradient self-weighting linear collaborative discriminant regression classification for human cognitive states classification. Mach Vis Appl. 2020;31:1–16.
21. Wen Z, Yu T, Yu Z, Li Y. Grouped sparse Bayesian learning for voxel selection in multivoxel pattern analysis of fMRI data. Neuroimage. 2019;184:417–30.
22. Kasabov NK. Deep Learning and Deep Knowledge Representation of fMRI Data. In: Time-Space Spiking Neural Networks and Brain-Inspired Artificial Intelligence. Berlin, Heidelberg: Springer; 2019. p. 361–95.
23. Wang, X., & Mitchell, T. (2002). Detecting cognitive states using machine learning. Iterim working paperWhiteld ML, Sherlock G, Saldanha A, Murray JI, Ball CA, Alexander KE, Matese JC, Perou CM, Hurt MM, Brown PO, Botstein.
24. Eddy W, Fitzgerald M, Genovese C, Lazar N, Mockus A, Welling J. The Challenge of functional magnetic resonance imaging. J Comput Graph Stat. 1999;8(3):545–58. https://doi.org/10.2307/1390875.
25. Hall, M. A. (1998) Correlation-based feature subset selection for machine learning. Thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy at the University of Waikato
26. Tsamardinos I, Borboudakis G, Katsogridakis P, Pratikakis P, Christophides V. A greedy feature selection algorithm for Big Data of high dimensionality. Mach Learn. 2019;108(2):149–202.
27. Kohavi R. Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid. In Kdd. 1996;96:202–7.
28. Rodriguez JJ, Kuncheva LI, Alonso CJ. Rotation forest: a new classifier ensemble method. IEEE Trans Pattern Anal Mach Intell. 2006;28(10):1619–30.
29. Shieh MD, Yang CC. Multiclass SVM-RFE for product form feature selection. Expert Syst Appl. 2008;35(1–2):531–41.
30. Behroozi M, Daliri MR. RDLPFC area of the brain encodes sentence polarity: a study using fMRI. Brain Imaging Behav. 2015;9(2):178–89.
31. Kasabov NK. NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. Neural Netw. 2014;52:62–76.
32. Doborjeh, M. G., Capecci, E., & Kasabov, N. (2014, December). Classification and segmentation of fMRI spatio-temporal brain data with a NeuCube evolving spiking neural network model. In 2014 IEEE Symposium on Evolving and Autonomous Learning Systems (EALS) (pp. 73–80). IEEE.
33. Ranjan, A., Singh, A. K., Thakur, A. K., Mishra, R. B., & Singh, V. P. (2021). Sentential Negation Identification of FMRI Data Using k-NN. In: Machine Intelligence and Smart Systems: Proceedings of MISS 2020 (pp. 657–664). Springer Singapore
34. Ranjan A, Singh VP, Singh AK, Thakur AK, Mishra RB. Classifying brain state in sentence polarity exposure: An ANN model for fMRI data. RIA. 2020. https://doi.org/10.18280/ria.340315.
35. Ranjan A, Singh VP, Mishra RB, Thakur AK, Singh AK. Sentence polarity detection using stepwise greedy correlation based feature selection and random forests: an fMRI study. Journal of Neurolinguistics. 2021;59:100985.