TABLE II. COMPARISON OF MAXIMUM CLASSIFICATION RESULTS

Subject (S)	Periodogram			Lomb-Scargle			Thompson Multitaper			Welch		
	MLNN	QSVM	KNN	MLNN	QSVM	KNN	MLNN	QSVM	KNN	MLNN	QSVM	KNN
S1	92.6	75.8	75.1	88.7	75.6	75.2	88.2	76.3	77.1	94.3	78.6	75.5
S2	86.4	78.6	78.8	90.8	79.5	78.1	92.7	78.3	79.2	91.9	81.7	80.2
S3	94.0	77.1	85.5	93.3	76.1	85.6	92.1	76.1	86.7	92.8	75.2	82.1
S4	86.9	70.8	72.1	92.2	69.6	73.2	86.9	68.6	72.0	90.6	69.2	72.0
S5	89.4	74.3	73.6	87.2	74.8	72.5	89.0	72.7	72.6	88.0	68.4	72.1
S6	84.0	62.2	68.4	85.3	63.7	68.2	88.2	63.0	66.6	82.0	**62.2**	**62.2**
S7	85.6	74.3	75.2	87.8	74.3	75.5	83.7	72.9	73.7	82.3	70.4	76.7
S8	93.3	74.1	79.4	**96.3**	72.1	78.1	94.1	73.7	79.1	90.2	76.4	73.6
S9	84.0	68.8	68.8	85.0	69.6	70.1	87.4	70.8	69.1	87.9	66.1	66.7
S10	89.3	65.3	74.9	92.0	65.9	73.9	89.0	67.2	75.8	87.1	67.6	76.1
Min (%)	84.0	62.2	68.4	85.0	63.7	68.2	83.7	63.0	66.6	82.0	**62.2**	**62.2**
SD	5.0	8.2	8.6	5.7	7.9	8.7	5.2	7.6	10.1	6.2	9.7	9.9
Max (%)	94.0	78.6	85.5	**96.3**	79.5	85.6	94.1	78.3	86.7	94.3	81.7	82.1

2) Quadratic SVM (QSVM)

Support Vector Machine (SVM) is developed by Boser, Guyon, and Vapnik [19] in 1992. The SVM can be used for regression and classification of the given input data into two or more classes [20]–[22]. The SVM performs better than MLNN in terms of training, generalization, no local optimal and scales relatively well to a large scale of dimension [23]. The SVM is particularly powerful and universal learning machine because of its characteristics such as ease of changing the implemented decision surface and capacity control obtained by optimizing the margin.

Basically, SVM works by finding a hyperplane, as in (5), that separates the positive and negative values of a given set of training vectors (dataset), as in (6), from each other with maximum margin to segregate the output classes, i.e. Alert and Fatigue.

$$\langle w, x \rangle + b = 0 \tag{5}$$

$$\text{Ɗ} = \{x_j, y_j\}_{j=1}^{N},\ x_j \in R^n,\ y_j \in \{-1,1\} \tag{6}$$

If the set of training vectors is separated without miscalculation and the distance between the nearest vector to the hyperplane is maximal, then it is said to be optimally separated by the hyperplane. However, in this works, for a nonlinearly separable data, a quadratic SVM is used to create an optimal hyperplane in a high-dimensional space with maximum margin to perform pattern recognition for a given set of training vectors to segregate the output classes. The general degree-two polynomial kernel is defined as in (7).

$$K(x, x') = (\langle x, x' \rangle + 1)^2 \tag{7}$$

where x and x' are vectors in the input space, which the polynomial kernel maps a two-dimensional input vectors into a six-dimensional feature space [22].

3) K-Nearest Neighbor (KNN)

KNN is a non-parametric classifier, supervised learning algorithm, and widely used for pattern classification in EEG studies [24], [25]. In pattern classification, the KNN classifies the feature samples using the nearest training pattern in the feature vectors and the target class is classified based on majority voting of its neighbors. To the extent that, the target class is associated with the most predominant class amongst its KNN measured by a distance function. The KNN classifier significantly depends on the value of k (positive integer) and the distance function used. Numerous methods have been done to optimize the value of k and different distance functions such as Euclidean, Manhattan, Minkowski, and Hamming.

In this study, the KNN algorithm is implemented using the Euclidean distance function to locate the nearest neighbor [26]. The Euclidean distance function measures $ED(x, y)$ between two samples (points) x and y using (8). The number of neighbor's 'K' is used to classify the new test vector was varied from 1 to 10, to determine the classification accuracy.

$$ED(x, y) = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} \tag{8}$$

IV. RESULTS AND CONCLUSION

From Table II, it can be perceived that the performance of MLNN classifier has the maximum classification rate of 96.3% using Lomb-Scargle PSD features set and minimum classification rate of 82.0% using Welch PSD features set. The performance of QSVM classifier has the maximum classification accuracy of 81.7% using Welch PSD features set and minimum classification accuracy of 62.2% using periodogram and Welch PSD features set. The performance of KNN classifier has the maximum classification correctness of 86.7% using Thompson multitaper PSD features set and minimum classification correctness of 62.2% using Welch PSD features set.

From the results, it is surmised that the performance of MLNN classifier model has the maximum classification accuracy of **96.3%** using Lomb-Scargle PSD features set and the performance of QSVM and KNN both have the minimum accuracy of **62.2%** using periodogram and Welch PSD features set respectively.