# Evaluation of Four Classification Techniques for X-ray Data Classification to Detect Caries

**Introduction:** The purpose of this report is to evaluate four classification techniques for the classification of Dental X-ray data to detect caries. The four techniques evaluated are decision trees, convolution neural networks (CNN), KNN, and support vector machines (SVM).

**Classifiers Evaluated:** The four classification techniques used in this work are briefly reviewed as follows:

1. Decision Trees: A statistical model for classification and data prediction that decomposes complex problems into simpler sub-models recursively. The C4.5 algorithm, an extension of the ID3 algorithm, was used for this study.
2. Convolution Neural Networks: A Deep Learning algorithm that assigns importance to various aspects/objects in an image and differentiates one from the other. The CNN architecture is analogous to the connectivity pattern of neurons in the human brain and is inspired by the organization of the visual cortex.
3. KNN: A classical prototype-based classifier that is often used in real-world applications due to its simplicity.
4. Support Vector Machines: A recent technique for classification and regression that has achieved remarkable accuracy in several important problems. SVM is based on the principle of structural risk minimization (SRM).

**Experiments:** The simulations were conducted using the Scikit-Learn and Keras libraries with TensorFlow backend in Python 3. The study used 10-fold cross-validation to assess the generalization performance and compare the classifiers. Stratified CV was used to form subsets with the same frequency distribution of patterns of the original data.

**Performance Measures:** The performance measures used to compare the classifiers were (1) the classification error and (2) the area under the ROC curve (AUC). The AUC summarizes the ROC curve and is another way to compare classifiers other than accuracy.

**Results and Discussion:** The experiment showed a huge imbalance in classes. The results demonstrated that the CNN achieved the best classification performance, with an AUC of 0.53. SVM had the second-best performance. KNN and decision trees had the lowest performance, respectively.

**Conclusion:** In conclusion, the study demonstrated that CNN and SVM were the most effective classification techniques for the classification of X-ray data to detect caries. The findings suggest that Deep Learning techniques may provide better classification results for this application than classical machine learning techniques.

| Model | Accuracy | Precision | Recall | F1 score | AUC |
|---|---|---|---|---|---|
| SVM | 51.13% | 23.44% | 48.39% | 31.58% | 51.64% |
| Non-linear SVM + PCA=2 | 55.64% | 15.62% | 66.67% | 25.32% | 47.71% |
| KNN | 56.39% | 12.50% | 80.00% | 21.62% | 52.46% |
| KNN + PCA=2 | 53.38% | 10.94% | 58.33% | 18.42% | 47.2% |
| Decision tree | 49.62% | 23.44% | 45.45% | 30.93% | 49.1% |
| Decision tree+depth=5 | 51.13% | 25.00% | 48.48% | 32.99% | 53.02% |
| Decision tree+pca=2 | 51.13% | 26.56% | 48.57% | 34.34% | 48.14% |
| Custom CNN | 53.38% | 75.36% | 53.61% | 62.65% | 48.1% |
| Inception Net | 48.12% | 78.26% | 50.00% | 61.02% | 46.24% |