IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.151148

EMPLOYMENT STATUS CLASSIFICATION USING SUPPORT VECTOR MACHINES OF VISUALLY-IMPAIRED GRADUATES

Neelesh N Shrinivisan[1], Aparna M[2], Hasanthika Sri S[3], G. Paavai Anand[4]

Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani[1,2,3]

Assistant Professor (S.G), Department of Computer Science and Engineering,

SRM Institute of Science and Technology, Vadapalani[4]

Abstract: This study uses the 2015-2019 American Community Survey (ACS) 5-year Public Use Microdata Sample (PUMS) for classifying the Employment Status of Visually-impaired Graduates. The original data had about 15.95 million records; it was then filtered to contain the population being visually impaired (or reported difficulty with vision), aged 21 to 64, and had at least a Bachelor's degree. Support Vector Machines (SVMs) are a supervised machine learning algorithm used for classification and regression tasks. It identifies the best boundary known as a hyperplane that separates different classes in the data. In this study, linear, radial basis function (RBF), and sigmoid kernels were compared using One-vs-One (OvO) and One-vs-Rest (OvR). The filtered dataset contains 29,714 rows and 23 columns. From the dataset fields, WKW, and WKHP a three-class target was derived. After the target-related fields were excluded, 14 predictors remained. The data was split into training and testing data in a 80:20 split and feature scaling was fitted on the training data. Each model was evaluated with and without balanced class weights. The highest reported accuracy was 65.02% for the unweighted RBF-OvO model, closely followed by RBF-OvR at 65.00%. Although the unweighted models achieved higher accuracy, they showed very low recall for the "Less than FT/FY" class. Balanced class weights improved recognition of the less frequent class in several configurations, while reducing overall accuracy. These observations show that accuracy alone is not enough to determine the performance of this model on the dataset.

Keywords: Support Vector Machines, Employment Status Classification, SVM Kernels, One-vs-Rest, One-vs-One, Class Imbalance

I. INTRODUCTION

In classification, a model is trained to assign observations to predefined categories. Its performance is evaluated by comparing predictions with actual labels. However, overall accuracy may not show how well the model performs for each category, especially when some classes have more observations than others.

Support Vector Machines (SVMs) are supervised learning methods that find a decision boundary with the largest possible margin between classes. Different kernels allow SVMs to handle different patterns. A linear kernel creates a straight boundary, while Radial Basis Function (RBF) and sigmoid kernels can represent nonlinear patterns. For problems involving more than two classes, SVMs can use One-vs-Rest (OvR) or One-vs-One (OvO) strategies [1], [2].

Employment-status classification was used to compare these approaches. Three target categories were derived from work-related fields in the supplied dataset. Since the classes had unequal numbers of observations, balanced class weights were also tested to give greater importance to less frequent classes during training.

Linear, RBF, and sigmoid kernels were compared using OvR and OvO, with and without balanced class weights. The study examined which configuration achieved the highest accuracy, how class weighting affected recognition of the less frequent class, and why class-wise measures were needed alongside accuracy. The work focused on comparing rather than validating a system for employment-related decisions.

II. RELATED WORK

Cortes and Vapnik established the foundation for modern Support Vector Machines (SVMs) through their work on support-vector networks, which introduced the use of large-margin decision boundaries for classification [1]. Kernel

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.151148

methods extend this approach by allowing models to identify similarities between observations without directly calculating every coordinate in a transformed feature space.

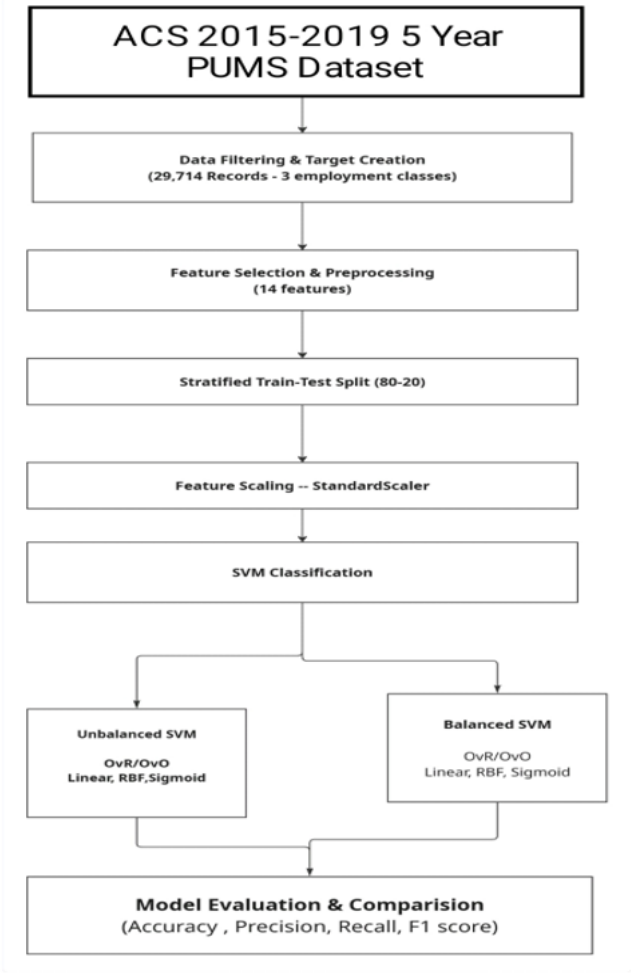
McDonnall et al., studied the factors that influence employment status and earnings amongst graduates with visual impairments using the American Community Survey (ACS) 2015-2019 5 Year Public Use Microdata Sample (PUMS) dataset. The study applied multinomial logistic regression to analyse the three classes. Their study indicated that demographic and socioeconomic factors are stronger predictors of employment and earning than the degree majors. [4] SVM models can be implemented using libraries such as Scikit-learn and LIBSVM. Scikit-learn provides the Support Vector classifier (SVC) interface for Python, while LIBSVM is widely used for support vector classification [2], [3]. These libraries allow different kernels and multiclass strategies to be tested under consistent data and evaluation conditions.

Class imbalance can affect how a model views different categories. The model may favour the majority class when one class has more observations than the others. As a result of this imbalance the model may achieve a good accuracy while missing less frequent classes. Evaluation metrics such as precision recall, F1-score helps reveal these differences. Macro-averaged F1 treats all classes equally, whereas weighted F1 gives more importance to classes with more observations. This study compares different kernels and multiclass strategies using the same data split, both with and without balanced class weights. Rather than testing every possible SVM configuration, it focuses on the results obtained using the preprocessing steps and model parameters specified.

III. MATERIALS AND METHODS

A. Flowchart/ Workflow

The overall workflow of the experiment is shown in Fig1:



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.151148

In this we began by using the data set of ACS 2015-2019 5-Year PUMS dataset. The next step is to filter the dataset to contain only the population with eye difficulty that is DEYE=1, ages 21 to 64, and SCHL (Educational attainment codes) within 21 to 24, resulting in a total of 29,714 records. Three employment classes were created using WKW(Weeks worked in the past 12 months) and WKHP(Hours worked per week in the past 12 months). After feature selection and pre-processing,we are left with 14 features. Then the models were trained on a 80-20 train-test split. OvO and OvR SVM models were trained on the same dataset and were compared with 3 kernels: Linear, RBF(Radial-Basis Function), Sigmoid. The results were compared between balanced and unbalanced class weights by using metrics such as Accuracy, Precision, Recall, F1 score.

The individual steps and stages are described in the following subsections.

B. Dataset and Target Definition

The American-Community Survey (ACS) Public Use Microdata Sample (PUMS), was used. The dataset altogether contained about 15.95 million population. Then it was filtered to contain the visually impaired people (or those who reported difficulty with vision), aged 21 to 64 and those graduates with at least a Bachelor's degree (SCHL in 21 to 24 range).[5]

The final filtered dataset contains 29,714 rows and 23 columns. The target variable, the class of the employment, was constructed from the WKW and WKHP fields. The target record was created and labelled “Not Employed” when WKW was missing. It was labelled “Full-Time /Full-Year (FT/FY) Employed” when WKW equalled 1 and WKHP was at least 35. All remaining records were assigned to “Less than FT/FY.”

This rule provides a consistent way for labels to be assigned in the experiment. This classification was based on the study by McDonnall et. al where the status was categorised as employed full-time/full-year (FT/FY) if they worked for at least 35 hrs per week and 50 to 52 in the previous year (past 12 months).[4]

The class-wise distribution of the classes is given below in Fig2.

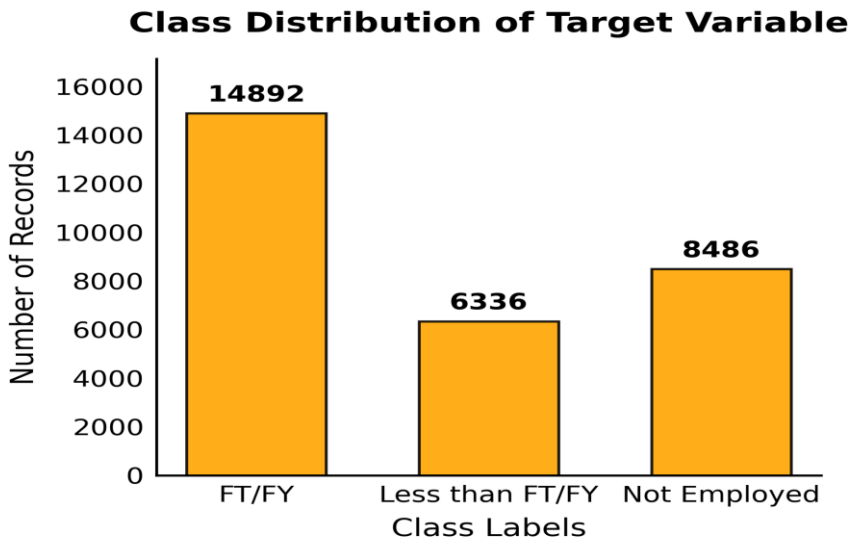


Fig.2. Class distribution of target variables(Employment_Status)

TABLE I. DISTRIBUTION OF THE DERIVED TARGET CLASSES

Target class	Records	Share
FT/FY	14,892	50.12%
Not Employed	8,486	28.56%
Less than FT/FY	6,336	21.32%
Total	29,714	100.00%

The FT/FY category is the largest class, accounting for approximately half of the records. The “Less than FT/FY” category is the smallest, with about one fifth of the data. This difference in class support is relevant to the interpretation of overall accuracy and provides the reason for testing balanced class weights.



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026
DOI: 10.17148/IJARCCE.2026.151148

C. Feature Selection and Missing Values

The target column "Employment Status" and the columns PWGTP, WKHP, WKW, WKWN, WRK, ESR, DEYE(Difficulty in Vision), WKL, and DIS (Disability recode, whether Disability exists or not) were dropped from the predictors, leaving 14 input features. The DIS column and DEYE columns were dropped as they contained a value of 1 throughout the records.

The target is excluded by definition, while the removal of fields used to construct the label helps reduce direct target leakage. This does not guarantee that all remaining predictors are independent of the target; other variables may still be associated with the work-related fields used to derive it.

In the workflow, missing FOD2P values were replaced with zero. A total of 26,413 missing values were present in this field. Although this treatment is easy to reproduce, zero may also be a valid code; if so, missing and genuine zero values may have been combined. The results should therefore be interpreted with this preprocessing choice in mind. Alternative imputation methods and missingness indicators could be examined in a later study.

D. Train-Test Split and Scaling

An 80:20 train-test split was applied with static random state value, so for each run the split did not get modified, and be stratified. Through stratification, approximately the same class proportions were retained in both partitions. The test partition contained 5,943 records and was reserved for evaluation; the scaler and classifiers were fitted using the training partition.

Feature scaling was performed with StandardScaler. This step is important for SVMs because differences in feature ranges can influence the geometry of the separating boundary. The scaler was fitted on the training data and then applied to the test data. This procedure avoids allowing the test distribution to influence the fitted scaling parameters. The same split and scaling approach was retained for the model comparisons so that the results could be compared on a consistent basis.

E. Kernels and Multiclass Strategies

Three kernel types were examined. The linear kernel represents a linear decision boundary in the scaled feature space. The RBF kernel can model nonlinear relationships through a distance-based similarity function. The sigmoid kernel applies a sigmoid-shaped transformation of the inner product. Their performance can depend on the dataset and the parameter settings, so no kernel is assumed to be best in advance.

Support Vector Machines (SVMs) are a binary classifier, we use two types of it for multiclass classification. In One-vs-One (OvO) classification, as the naming suggests, it means One class vs one of the other classes. So, in this OvO, we get

1. FT/FY vs Less than
2. FT/FY vs Not Employed
3. Less than FT/FY vs Not Employed

Consider K=3 since its a 3 class classifier, the general formula of number of classifiers constructed is

$K(K-1)/2 = 3$ which gives us 3 classes.

One-vs-Rest (OvR) as the naming suggests, its One class vs the other remaining classes

1. FT/FY vs Not FT/FY
2. Less than FT/FY vs Not less than FT/FY
3. Not Employed vs Not Not Employed

Considering K = 3 as its a 3 class classification, the general formula of the number of classifiers constructed is equal to K, which equals 3 for a multi-class problem.

The two multiclass strategies were compared as explained above. The six unweighted configurations were linear-OvR, linear-OvO, RBF-OvR, RBF-OvO, sigmoid-OvR, and sigmoid-OvO.

The second experiment repeated these configurations with class_weight set to "balanced" for the unequal class distributions.

In this setting, class weights are adjusted inversely according to the class frequencies in the training data. This change was used to examine whether minority-class recognition could be improved. Apart from class weighting, the experiments

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.151148

did not use an extensive hyperparameter search; the results should therefore be treated as a comparison of the tested configurations rather than as optimized performance estimates.

F. Evaluation Measures

Accuracy is determined by the percentage of correctly classified test records. Precision indicates how many of the records which have been predicted to belong to a particular class are correctly classified. Recall shows how many of the actual records are identified correctly in that class. The F1-score is obtained by combining precision and recall through their harmonic mean. Macro F1 is calculated by giving equal importance to the F1-score of each class, whereas weighted F1 is calculated by considering the number of records in each class. All these measures were used together because a model may achieve good accuracy while still showing weaker performance on classes with fewer records.

IV. RESULTS

A. Overall Comparison

The reported results show that the unweighted RBF-OvO configuration achieved the highest accuracy at 65.02%, narrowly ahead of RBF-OvR at 65.00%. The linear configurations achieved accuracies of 64.26% for OvR and 63.96% for OvO. The sigmoid configurations were lower: 52.04% for OvR and 54.79% for OvO. The difference between the two RBF strategies was only 0.02%, so the available single-split results do not establish a meaningful practical advantage for one over the other.

TABLE II. ACCURACY WITH AND WITHOUT BALANCED CLASS WEIGHTS

Kernel / strategy	Unweighted accuracy	Balanced accuracy
Linear / OvR	64.26%	60.81%
Linear / OvO	63.96%	49.96%
RBF / OvR	65.00%	59.11%
RBF / OvO	65.02%	60.09%
Sigmoid / OvR	52.04%	42.07%
Sigmoid / OvO	54.79%	49.15%

The balanced runs had lower accuracy in all six configurations. The largest reported unweighted accuracy was therefore retained by RBF-OvO, but this result alone does not describe how the model handled each class. The reports show that the unweighted RBF-OvO model had a macro F1-score of approximately 0.48 and weighted F1 of approximately 0.57. The gap between the two scores reflects uneven performance across classes, rather than uniformly strong classification.

B. Class-Wise Behaviour

The class-wise reports reveal the most important limitation in the unweighted configurations: recall for “Less than FT/FY” was close to zero in the linear and RBF configurations. For RBF-OvO, only six records from this class were correctly identified in the confusion matrix. This indicates that the model rarely assigned the less frequent category, even though the overall accuracy was close to 65%. The result demonstrates why an aggregate measure should not be used as the only basis for selecting a classifier.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.151148

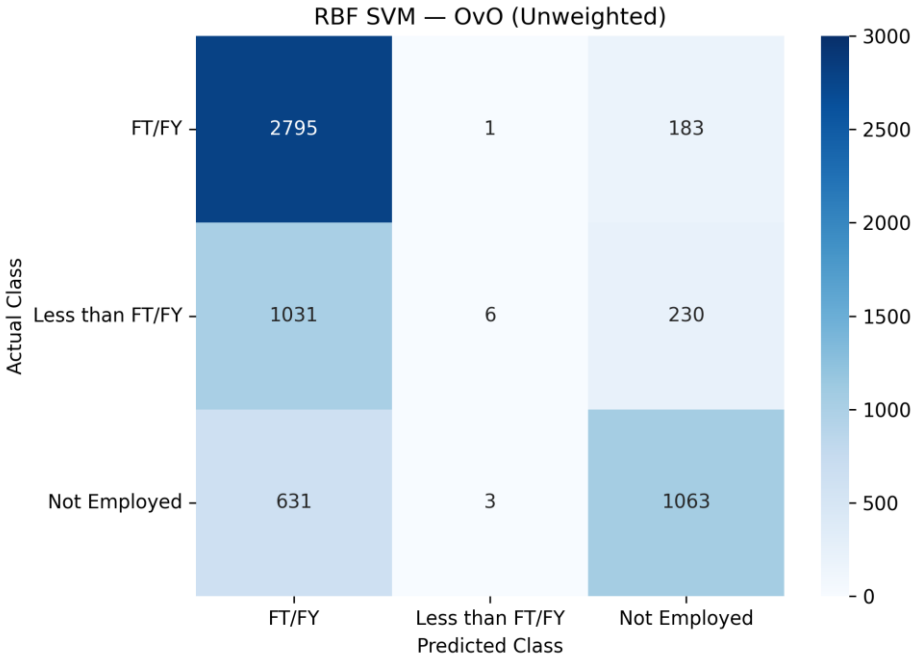


Fig 3. Confusion Matrix unbalanced OvO, RBF kernel

In the confusion matrix represented as heatmap in Fig3, the rows represent actual classes and columns represent predicted classes. Most FT/FY records were correctly assigned, while a substantial number of "Less than FT/FY" records were predicted as FT/FY. Most Not Employed records were correctly identified, although some were also assigned to FT/FY. The near-zero recall for "Less than FT/FY" is not proof that this category is inherently unpredictable. It may be related to the target rule, overlapping predictor patterns, class imbalance, missing-value treatment, or the selected model parameters.

The sigmoid models had lower overall accuracy but displayed a different class-wise pattern. In the saved reports, sigmoid-OvO reached a recall of approximately 0.15 for "Less than FT/FY," and sigmoid-OvR reached approximately 0.14. This was still limited recognition, but it was higher than the near-zero recall observed for several unweighted linear and RBF configurations. These differences reinforce the need to inspect the class-wise results rather than ranking the models by accuracy alone.

C. Effect of Balanced Class Weights

Balanced class weights changed the balance between overall accuracy and minority-class recognition. For RBF-OvR, accuracy decreased from 65.00% to 59.11%; for RBF-OvO, it decreased from 65.02% to 60.09%. Linear-OvO decreased from 63.96% to 49.96%, while sigmoid accuracy also declined. These reductions indicate that giving greater weight to less frequent classes changed the decision boundaries and increased errors elsewhere in the test set.

The main benefit was improved recall for the less frequent class in several configurations. For RBF-OvR, recall for "Less than FT/FY" increased from approximately zero to 0.33; for RBF-OvO as shown in Fig. 4 below, it increased to approximately 0.30. This is a substantial change in class recognition, although the gain came with lower overall accuracy. Hence, the choice between weighted and unweighted models depends on whether the application prioritizes total correct predictions or balanced performance across classes.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.151148

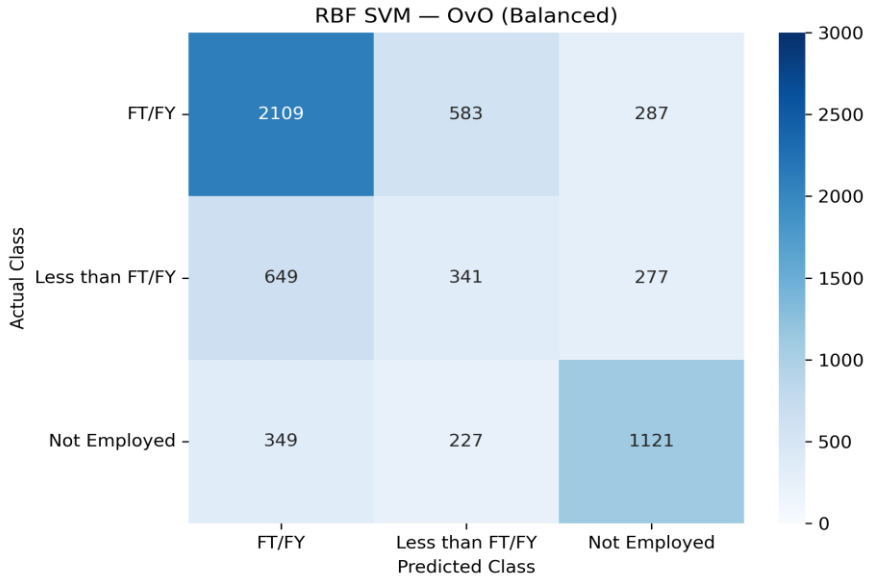


Fig 4. Confusion matrix for balanced RBF-OvO model

Class weighting should not be treated as a complete solution to class imbalance. It cannot create additional examples or correct possible weaknesses in the target definition. It can also increase false-positive predictions for a class when recall improves. A stronger study would compare class weighting with resampling and threshold adjustment inside a cross-validation pipeline, with all such operations restricted to training folds.

D. Evaluation Beyond Accuracy

Accuracy summarises the share of records classified correctly, but it does not show which classes account for the errors. This distinction is important in the present dataset because the FT/FY category represents about half of the records, while "Less than FT/FY" represents about one fifth. Moderate accuracy may still be achieved by a model when the larger category is predicted more often, even if only a few records from the smaller category are identified.

A more detailed evaluation is provided by precision, recall, and F1-score. Precision is measured by how many predictions for a class are correct, while recall is measured by how many actual records of that class are identified. The F1-score is calculated using both precision and recall. Equal importance is given to all classes in macro F1, while greater influence is given to classes with more records in weighted F1. For this reason, macro F1 is useful when performance across all classes is important, whereas weighted F1 describes performance with the observed class distribution taken into account.

The confusion matrix should be read together with these measures. In the unweighted RBF-OvO result, only six records from the "Less than FT/FY" class were correctly predicted, while 1,031 were classified as FT/FY and 230 as Not Employed. This pattern explains why the model's overall accuracy does not reflect balanced performance across categories. The counts describe the held-out partition only and should not be interpreted as evidence that the same error rates would occur in a different population.

V. DISCUSSION

A. Interpretation of the Main Findings

For the first research question, the highest unweighted accuracy was achieved by RBF-OvO, although the difference from RBF-OvR was very small. Since only one stratified hold-out split was used, this difference should not be considered enough to establish that one model is better than the other. Instead, it can be concluded that both RBF configurations performed similarly and achieved the highest overall accuracy among the unweighted models tested.

For the second research question, a decrease in accuracy was observed when balanced class weights were used. However, recall for the less frequent class was improved in several configurations. This is expected because greater importance is given to errors involving less frequent classes during model training. Whether this trade-off is useful depends on the purpose of the model and the importance given to different types of prediction errors. Since no application-specific cost matrix was provided, it cannot be determined from these results which trade-off would be more suitable for an actual decision-making process.



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026
DOI: 10.17148/IJARCCE.2026.151148

For the third research question, accuracy alone was found to be insufficient because correct predictions from all classes are combined into a single measure. In this study, almost no records from the "Less than FT/FY" class were identified by a model that achieved approximately 65% accuracy. This limitation was shown more clearly through macro F1 and per-class recall. Weighted F1 was also useful as an overall measure because the class distribution was taken into account. However, greater influence was given to the larger classes. Therefore, these measures should be considered together when model performance is evaluated.

B. Reproducibility and Experimental Controls

*T*he results can be reproduced only when the same CSV file, target construction, feature exclusions, random seed, data split, scaling method, and SVC settings are used. Through stratification, the possibility of a large difference in class distribution between the training and testing sets is reduced. The scaler was also fitted using only the training data, so information from the test data was not used during preprocessing. The same features and test partition were used for each model configuration to ensure that the results could be compared fairly.

An extensive search for parameters such as C and gamma was not carried out, and repeated cross-validation results were not included. For this reason, the differences between the model scores should be viewed as observations from the current experiment rather than as evidence of a consistent performance difference. In future work, repeated stratified cross-validation could be performed on the training data, with a measure such as macro F1 selected for parameter tuning. The final model could then be evaluated once using a separate test set. A majority-class baseline and other classifiers could also be included to determine how much improvement is provided by the SVM models.

C. Limitations

A specific rule was used to create the target label. Records with missing WKW values were classified as Not Employed. Records with WKW equal to 1 and WKHP of at least 35 were classified as FT/FY, while all remaining records were assigned to Less than FT/FY. Although this rule was used in the notebooks, it has not been independently verified as a complete definition of employment status. Any ambiguity in the target label is also carried over to the classifiers and may affect the errors produced by the models.

A major limitation of this target definition is that it does not incorporate the Employment Status Recode (ESR), as the classification was entirely based on past 12-month work activity (WKW and WKHP). [5]

The dataset was obtained from 2015-2019 ACS 5-Year PUMS, published by the U.S Census Bureau. Although, the survey methodology, dataset source, and collection period are documented. This study does not incorporate the survey weights. Therefore, the observed proportions and model performance should not be interpreted as though their estimates represent the complete population.

The number of predictors was reduced to avoid direct leakage from fields related to the target variable. However, some of the remaining features may still be related to the excluded variables and act as indirect proxies. The analysis is based on prediction rather than causation, so no causal relationship between individual features and employment status can be established from the classification results. In addition, the use of a single data split limits the assessment of how results may vary across different samples. Small differences in accuracy may not remain the same when a different split or parameter selection process is used.

Missing values in FOD2P were replaced with zero, even though a large number of entries were missing. If zero is also a valid category code, missing values and actual zero values may have been treated as the same, which might have imposed artificial numeric scale on a nominal categorical feature during standard scaling.

Privacy, fairness, and error rates across different subgroups were also not examined in this experiment. Since employment status is an important outcome, any real-world application would require a clearly defined purpose, legal and ethical review, privacy protection, evaluation across subgroups, and human oversight. Therefore, this study should be considered an educational comparison of models rather than a system ready for practical use.

D. Future Work

The future work could include Census variable ESR along with the weeks worked (WKW), so that people who currently do not have a job are clearly separated from the population that has worked part of the year. In addition, categorical columns should be converted using proper encoding rather than treating them as numbers.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026
DOI: 10.17148/IJARCCE.2026.151148

While this study focused on Support Vector Machines, it could be worked with tree-based algorithms like Random Forest, and XGBoost (Extreme Gradient Boosting). Future work should compare SVM performance against these models.

Instead of relying on a single-train test split, future work could use a stratified k-fold cross-validation on the training data. Scaling and any resampling should be performed separately within each training fold to prevent information leakage. A limited search over relevant SVM parameters could then be conducted using a performance measure selected before the models are compared. The results should be reported using accuracy, macro and weighted F1, class-wise precision and recall, confusion matrices, and uncertainty estimates. Baseline models should also be included so that the performance of the SVM configurations can be compared with simpler alternatives.

In future studies, differences in prediction errors across relevant subgroups could also be examined where appropriate and legally permitted. Care should be taken when reporting these results so that information about small groups is not exposed. A single fairness metric should not be treated as proof that a model is fair. Instead, the aim should be to identify possible limitations and risks before any real-world application is considered. Finally, incorporating person-level survey weights (PWGTP) provided by the Census Bureau will help ensure that model predictions and findings represent a wider population of visually impaired graduates accurately.

E. Implications for Responsible Use

Employment predictions may affect the opportunities and services available to individuals. Therefore, the findings of this study should not be viewed as a basis for making decisions about individuals through automated systems. The study does not consider whether consent was obtained, how the data was handled, or how privacy was maintained. It also does not look into the possible consequences of incorrect predictions or whether certain demographic or other relevant groups are more affected by these errors. These aspects would need to be examined before the model could be considered for use.

When evaluating a model, its intended use and the possible consequences of its errors should both be taken into account. In this study, the model is used to compare different algorithms and examine the effect of class imbalance for educational purposes. However, its use in an actual service would require further checks. The target labels would need to be validated, the data source properly documented, and the necessary permissions obtained. The model would also need to be evaluated across different groups, monitored regularly, and used with appropriate human oversight. Therefore, the findings of this study are limited to an educational comparison of models and should not be considered evidence of their suitability for practical use.

F. Statistical Caution and Use in Other Settings

The scores reported in this paper were obtained using a single test set. They show how the models performed on the records in that set, but they do not indicate how much the results might change when a different sample is used. No confidence intervals or significance tests were calculated in the supplied notebooks. Therefore, small differences between model configurations should be reported as differences observed in this experiment rather than as improvements. Repeated stratified cross-validation could help show how much the results vary across different training and validation folds. However, a separate independent test set would still be needed for the final evaluation.

The distribution of the target classes should also be kept in mind when interpreting the results. The proportions reported in this study are based on the supplied dataset after the target rule was applied. They should not be treated as estimates of the actual population proportions without information about the sampling method and any relevant survey design. Stratification was used to maintain similar class proportions in the training and testing sets. However, this does not correct any sampling bias that may already be present in the original dataset. It only helps maintain a similar class distribution across the two sets.

The metric used to select a model should be decided before the final tuning begins. If all three classes are considered equally important, macro F1 and per-class recall may provide a more suitable evaluation than accuracy alone. If different prediction errors have different consequences in the intended application, these should also be considered when choosing the evaluation metric or cost structure. No such application-specific priorities were provided in this study. Therefore, the results are used to compare the observed trade-offs without suggesting that one weighting condition is always better than the other.

The results obtained should be verified by running the experiments before the paper is submitted. If the saved results and the new results do not match, the dataset, software environment, preprocessing steps, and model settings should be checked to identify the fault. This is particularly important when reviewing the confusion matrix and evaluation metrics.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026
DOI: 10.17148/IJARCCE.2026.151148

A change in the order of the classes could affect how the confusion matrix is interpreted, even if the numerical values remain unchanged.

VI. CONCLUSION

Linear, RBF, and sigmoid SVM kernels were compared using the One-vs-Rest and One-vs-One strategies for three-class employment-status classification. Among the unweighted configurations tested, RBF-OvO achieved the highest accuracy of 65.02%, followed closely by RBF-OvR at 65.00%. The sigmoid configurations showed lower accuracy in this experiment. However, these results are specific to the supplied dataset, preprocessing steps, data split, and model settings. They do not show that one SVM kernel will always perform better than the others.

The main finding of this study is related to how class imbalance affected the predictions for individual classes. Several unweighted configurations performed poorly in identifying the "Less than FT/FY" category, even though their overall accuracy was moderate. The use of balanced class weights improved recall for the less frequent class in several configurations, but this also led to a reduction in overall accuracy. This shows that model selection should depend on the intended objective and that accuracy alone may not provide a complete picture of performance. Class-wise measures and confusion matrices should also be considered.

There are some limitations to the findings of this study. The models were also not evaluated through extensive parameter tuning or repeated cross-validation. Future work should focus on checking the accuracy of the target labels, providing clearer information about the data source, and comparing the models with stronger baseline methods. It would also be useful to examine how consistent the model errors are and consider any ethical concerns. Until these issues are addressed, the results should be treated as a comparison of classification methods rather than proof that the models can be used for employment-related decisions.

REFERENCES

[1]. C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, pp. 273–297, 1995, doi: 10.1007/BF00994018.
[2]. C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, Article 27, 2011, doi: 10.1145/1961189.1961199.
[3]. F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[4]. M. C. McDonnall, J. L. Cmar, and Z. McKnight, “How Degree Major and Demographic Factors Influence Employment and Earnings for College Graduates with Visual Impairments,” *Rehabilitation Counseling Bulletin*, vol. 68, no. 1, pp. 13–24, 2024, doi: 10.1177/00343552231187587.
[5]. U.S. Census Bureau. 2015-2019 American Community Survey 5-year Public Use Microdata Samples(PUMS) Data Dictionary. Retrieved from https://www2.census.gov/programs-surveys/acs/tech_docs/pums/data_dict/PUMS_Data_Dictionary_2015-2019.pdf

 This work is licensed under a Creative Commons Attribution 4.0 International License