# **EXPERIMENT REPORT**

|  |  |
| --- | --- |
| **Student Name** | Aibarna Singh Basnet |
| **Project Name** | AT1 |
| **Date** | 31/08/2023 |
| **Deliverables** | basnet\_aibarna-24585717-week3\_best\_model\_selection.ipynb  https://github.com/aibarna96/adv\_mla\_at1 |

|  |  |
| --- | --- |
| 1. **EXPERIMENT BACKGROUND** | |
| Provide information about the problem/project such as the scope, the overall objective, expectations. Lay down the goal of this experiment and what are the insights, answers you want to gain or level of performance you are expecting to reach. | |
| **1.a. Business Objective** | Explain clearly what is the goal of this project for the business. How will the results be used? What will be the impact of accurate or incorrect results?  The goal of this project is to develop a predictive model that classifies basketball players as potential draft picks for the upcoming season based on specific attributes. The intended business outcome is to aid basketball teams in optimizing their player recruitment process by identifying prospects with higher draft potential. The model's accurate predictions will assist teams in efficiently allocating resources, forming strategically balanced rosters, and gaining a competitive edge. However, incorrect predictions could lead to misallocation of resources and missed opportunities for team improvement. The project's success lies in providing actionable insights for effective player selection, thereby enhancing team performance and overall success. |
| **1.b. Hypothesis** | Present the hypothesis you want to test, the question you want to answer or the insight you are seeking. Explain the reasons why you think it is worthwhile considering it,  We aim to investigate whether specific player attributes, such as scoring efficiency, defensive performance, and overall productivity, have a substantial impact on a basketball player's likelihood of being drafted in the next season. Our hypothesis suggests that these attributes play a crucial role in draft prospects. By exploring the correlations and predictive power of these attributes, we aim to uncover insights that could enhance player recruitment strategies. This investigation is worthwhile as accurate predictions could optimize resource allocation and inform recruitment decisions, ultimately contributing to team success. |
| **1.c. Experiment Objective** | Detail what will be the expected outcome of the experiment. If possible, estimate the goal you are expecting. List the possible scenarios resulting from this experiment.  In my experiment, I assessed different classification algorithms to predict whether a college basketball player will be drafted to the NBA based on their current season statistics. I used AUROC as the performance metric to evaluate the models. The goal was to find the best-performing model with a high AUROC score on the test dataset. The RandomForestClassifier achieved the highest AUROC score of 0.9975, making it the top choice for predicting player drafts. However, I need to ensure the model doesn't overfit, consider model interpretability, and address potential data limitations. External validation and hyperparameter tuning are also important steps to validate and fine-tune the model's performance. Ultimately, the aim is to build a reliable model that effectively distinguishes between drafted and non-drafted players in the NBA draft process. |

|  |  |
| --- | --- |
| 1. **EXPERIMENT DETAILS** | |
| Elaborate on the approach taken for this experiment. List the different steps/techniques used and explain the rationale for choosing them. | |
| **2.a. Data Preparation** | Describe the steps taken for preparing the data (if any). Explain the rationale why you had to perform these steps. List also the steps you decided to not execute and the reasoning behind it. Highlight any step that may potentially be important for future experiments  For data preparation, we began by loading the training dataset and transforming column names to lowercase for consistency. We then selected relevant features with strong correlations to the target variable "drafted." As missing values were present, we filled them with zeros for simplicity. The data was split into features and the target variable, followed by a division into training and testing sets. To address class imbalance, we employed SMOTE to oversample the minority class. We standardized the features using StandardScaler to ensure model effectiveness. While we treated missing values by filling with zeros, other advanced techniques like imputing could be explored for improved data quality. Additionally, feature engineering, interactions, or polynomial features might enhance model performance. The rationale for these steps is to ensure data is suitable for model training and that the model is provided with relevant information. |
| **2.b. Feature Engineering** | Describe the steps taken for generating features (if any). Explain the rationale why you had to perform these steps. List also the feature you decided to remove and the reasoning behind it. Highlight any feature that may potentially be important for future experiments  For feature generation, we started by selecting attributes that we hypothesized to be relevant indicators of a player's draft potential, including scoring efficiency, defensive performance, and overall productivity metrics. We combined features such as "porpag" (points per offensive rebound), "dunksmade" (number of successful dunks), "dporpag" (defensive rebounds per offensive rebound), and others. We also engineered "midmade\_midmiss," which represents the ratio of successful mid-range shots to unsuccessful ones. To enhance interpretability, we normalized features using StandardScaler. However, we noticed multicollinearity between "dunksmade" and "dunksmiss\_dunksmade," leading us to remove the latter due to redundancy. Notably, "midmade" and "midmade\_midmiss" could be vital for future experiments as they might reflect a player's proficiency in mid-range shooting, an area of increasing importance in modern basketball strategy. |
| **2.c. Modelling** | Describe the model(s) trained for this experiment and why you choose them. List the hyperparameter tuned and the values tested and also the rationale why you choose them. List also the models you decided to not train and the reasoning behind it. Highlight any model or hyperparameter that may potentially be important for future experiments  In this experiment, I trained four classification models to predict whether a college basketball player will be drafted to the NBA based on their statistics. The models I chose were Logistic Regression, RandomForestClassifier, XGBClassifier, and GradientBoostingClassifier. I selected these models because they represent a mix of simple and complex algorithms, allowing me to gauge both interpretability and predictive power. For hyperparameter tuning, I focused on key parameters such as learning rate, number of estimators (trees), maximum depth, and minimum samples per leaf for the tree-based models. Specifically, for RandomForestClassifier, I tuned the 'n\_estimators' and 'max\_depth' parameters, while for XGBClassifier and GradientBoostingClassifier, I tuned 'learning\_rate', 'n\_estimators', 'max\_depth', and 'min\_samples\_leaf'. I chose a range of values for each hyperparameter to find the combination that maximizes performance, taking into account the trade-off between model complexity and generalization. I decided not to train more complex models like Neural Networks due to their potential overfitting risks given the relatively small dataset. Additionally, I didn't explore more unconventional algorithms due to their complexity and the need for specialized tuning. A key takeaway for future experiments could be the importance of fine-tuning the learning rate, as this hyperparameter significantly affects convergence speed and overall model performance for gradient-boosting-based algorithms. Also, experimenting with feature engineering and addressing class imbalance could be crucial for improved predictive power in future iterations of this experiment. |

|  |  |
| --- | --- |
| 1. **EXPERIMENT RESULTS** | |
| Analyse in detail the results achieved from this experiment from a technical and business perspective. Not only report performance metrics results but also any interpretation on model features, incorrect results, risks identified. | |
| **3.a. Technical Performance** | Score of the relevant performance metric(s). Provide analysis on the main underperforming cases/observations and potential root causes.  In my classification experiment, I evaluated multiple algorithms using AUROC as the primary performance metric. The RandomForestClassifier performed exceptionally well, achieving a test AUROC score of 0.9975, which is indicative of its strong predictive power in determining whether a college basketball player will be drafted to the NBA. However, during the analysis, I noticed a few observations where the model underperformed. These cases might be attributed to various factors such as overfitting, imbalanced data, or potential outliers that are influencing the predictions. It's important to carefully examine these underperforming cases, possibly through visualization and feature analysis, to identify potential root causes and take corrective measures. This could involve re-evaluating the model's complexity, applying techniques to balance the class distribution, or addressing outliers in the data. Ensuring that the model generalizes well to various scenarios is crucial for its real-world applicability. |
| **3.b. Business Impact** | Interpret the results of the experiments related to the business objective set earlier. Estimate the impacts of the incorrect results for the business (some results may have more impact compared to others)  Interpreting the results of my classification experiments is crucial in understanding their implications for the business objective of predicting whether a college basketball player will be drafted to the NBA. After evaluating multiple classification algorithms using AUROC as the metric, I found that the RandomForestClassifier performed the best with a high AUROC score of 0.9975 on the test data. This suggests that the model has a strong ability to differentiate between drafted and non-drafted players based on their current season statistics. This outcome aligns with my goal of selecting a model with high predictive accuracy.  However, it's important to consider the impacts of incorrect predictions for the business. If the model incorrectly predicts that a talented player will not be drafted, this could result in missed opportunities for both the player and the team. On the other hand, if the model incorrectly predicts that a player will be drafted when they are not, this could lead to misguided investments and team decisions. The potential impact of false negatives (missed drafted players) could be substantial, as valuable talents might be overlooked. While false positives (incorrectly predicted drafted players) could lead to wasted resources, their impact might be comparatively less severe. Therefore, ensuring a balance between precision and recall is important in this context to minimize both types of errors and maximize the model's overall usefulness for NBA team scouting.  In summary, the experiment's results show a promising model for predicting NBA draft selections based on player statistics. However, understanding the potential consequences of incorrect predictions is crucial for the business's decision-making process. Striking the right balance between minimizing false negatives and false positives is essential to maximize the model's impact and value for NBA teams. |
| **3.c. Encountered Issues** | List all the issues you faced during the experiments (solved and unsolved). Present solutions or workarounds for overcoming them. Highlight also the issues that may have to be dealt with in future experiments.  During the experiments, I encountered a few challenges, both solved and unsolved. One issue was the potential of overfitting, where some models showed very high performance on the training set but lower performance on the test set. To address this, I performed hyperparameter tuning and cross-validation to find the optimal model complexity that balances training and generalization. Another challenge was class imbalance in the dataset, which could bias the model's predictions. I addressed this by exploring techniques like oversampling and undersampling to balance the classes. Furthermore, I prioritized model performance based solely on AUROC scores, which might not capture the entire picture of a model's effectiveness. In future experiments, I plan to consider a wider range of evaluation metrics, including precision, recall, and F1-score, to gain a more comprehensive understanding of the model's performance. Additionally, I should assess the impact of external validation and explore potential biases in the data that could affect the model's fairness and generalization capabilities. Finally, I will aim to understand the interpretability of the chosen model, as a balance between performance and interpretability is crucial for practical applications. |

|  |  |
| --- | --- |
| 1. **FUTURE EXPERIMENT** | |
| Reflect on the experiment and highlight the key information/insights you gained from it that are valuable for the overall project objectives from a technical and business perspective. | |
| **4.a. Key Learning** | Reflect on the outcome of the experiment and list the new insights you gained from it. Provide rationale for pursuing more experimentation with the current approach or call out if you think it is a dead end.  Based on the results of the experiment, I gained valuable insights into the performance of different classification algorithms for predicting NBA draft selection based on player statistics. It's clear that the RandomForestClassifier outperformed other models, achieving the highest AUROC score of 0.9975 on the test dataset. This suggests that the model has a strong ability to distinguish between drafted and non-drafted players. However, I also learned that achieving extremely high AUROC scores might raise concerns about potential overfitting to the training data. Moreover, while a complex model like RandomForestClassifier performs well, it might lack the interpretability that simpler models offer. Considering this, further experimentation could involve investigating the potential for model overfitting, fine-tuning hyperparameters to ensure robustness, exploring the trade-off between model complexity and interpretability, and conducting external validation on unseen data to truly evaluate generalization. This experimentation could lead to a more refined model that strikes the right balance between performance and interpretability. |
| **4.b. Suggestions / Recommendations** | Given the results achieved and the overall objective of the project, list the potential next steps and experiments. For each of them assess the expected uplift or gains and rank them accordingly. If the experiment achieved the required outcome for the business, recommend the steps to deploy this solution into production.  Based on the results achieved and the main goal of predicting NBA draft selection using classification algorithms, here are potential next steps and experiments to consider:  1. Hyperparameter Tuning: Perform a more exhaustive hyperparameter search for the RandomForestClassifier to fine-tune the model's parameters. This could potentially lead to improvements in model performance. Expected uplift: Moderate. Ranking: High priority.  2. Ensemble Methods: Experiment with ensemble methods such as stacking or boosting, combining the strengths of multiple models. This might further enhance predictive accuracy. Expected uplift: Moderate. Ranking: Medium priority.  3. Feature Engineering: Explore feature engineering techniques to create new informative features from the existing ones. This might uncover hidden patterns in the data and contribute to better predictions. Expected uplift: Moderate. Ranking: Medium priority.  4. External Validation: Validate the selected model on an external dataset that it has never seen before. This provides a more robust assessment of the model's generalization ability. Expected uplift: Validation of current performance. Ranking: High priority.  5. Explainability Analysis: Conduct explainability analysis to understand how the model arrives at its predictions. This can build trust and provide insights into players' characteristics influencing draft decisions. Expected uplift: Improved model interpretability. Ranking: Medium priority.  6. Deployment and Monitoring: If the selected model performs well on external validation and meets the business requirements, prepare to deploy it in a production environment. Monitor its performance and recalibrate periodically. Expected uplift: Real-world application. Ranking: High priority.  In conclusion, considering the high AUROC scores achieved and the goal of predicting NBA draft selections, the next steps should focus on refining the chosen RandomForestClassifier model through hyperparameter tuning and exploring ensemble methods. Additionally, validating the model on an external dataset, enhancing model interpretability, and preparing for deployment are crucial aspects. If the model's performance aligns with business requirements, deploy it in a production environment and establish monitoring practices to ensure its continued effectiveness. |