



Neapolis University Pafos

**Predicting NBA Rookie Impact from NCAA Statistics: A
Machine Learning Approach Using Regularized Adjusted
Plus-Minus (RAPM)**

**Authored by
Maksim Misevich**

Supervised by Dr. Andrei Smolensky

May 2026



Neapolis University Pafos

**Predicting NBA Rookie Impact from NCAA Statistics: A
Machine Learning Approach Using Regularized Adjusted
Plus-Minus (RAPM)**

**Dissertation which was submitted for obtaining an Applied
Computer Science degree**

**Authored by
Maksim Misevich**

**Supervised by Dr. Andrei Smolensky
May 2026**

Copyrights

Copyright © Maksim Misevich, 2026. All rights reserved.

The approval of the Dissertation by Neapolis University does not necessarily imply the acceptance of the author's views on behalf of the University

Table of Contents

Table of Contents	4
Abstract	7
List of Figures	8
List of Tables	9
Chapter 1 Introduction	10
Chapter 2 Literature Review	10
2.1. Advances in Measuring Individual Player Impact	10
2.2. NCAA-to-NBA Modelling Research.....	11
2.3. Key Limitations and Research Gaps.....	12
2.4. Contributions from the Present Study	12
Chapter 3 Data	13
3.1. Data Sources.....	13
3.2. Dataset Construction	13
3.3. Target Variable: Rookie Window RAPM	14
3.4. Feature Engineering	16
3.4.1 Adjusting For Different Eras	17
3.4.2 Possession Estimation and Per 100 Possessions Features	18
3.4.3 Team Context Features	20
3.4.4 Draft Age	21
3.5. Data Pipeline Flowchart.....	23
Chapter 4 Methodology.....	25
4.1. Problem Framing	25
4.2. Train/Test split	25
4.3. Feature Selection	25
4.3.1 Domain-Based Exclusions	25
4.3.2 Feature Redundancy Analysis.....	26
4.4. Modelling	27
4.4.1 Baseline Model Training and Hyperparameter Tuning	27
4.4.2 Feature Set Search	28
4.4.3 Feature Importance Comparison.....	29
4.4.4 Comparison Against Pre-Draft Consensus	34
Chapter 5 Discussion.....	35
5.1. Interpretation of Results.....	35
5.2. Cross-Model Agreement.....	35
5.3. Comparison to the Literature	35

5.4. Practical Applications.....	36
5.5. Limitations	36
Chapter 6 Conclusion	37
Chapter 7 References	37
Appendix.....	39

Validity Page

Student's Name & Surname: Maksim Misevich

Dissertation Title: Predicting NBA Rookie Impact from NCAA Statistics: A Machine Learning Approach Using Regularized Adjusted Plus-Minus (RAPM)

This Dissertation was prepared in the context of the studies for obtaining a Degree at Neapolis University Pafos and was approved on 22 May 2026 by the members of the Examiners' Committee.

Examiners' Committee:

First Supervisor (Neapolis University Pafos): Dr. Andrei Smolensky

Committee member: Dr. Avgousta Kyriakidou-Zacharoudiou

DECLARATION

I, Maksim Misevich, being fully aware of the consequences of plagiarism, declare responsibly that this paper entitled "Predicting NBA Rookie Impact from NCAA Statistics: A Machine Learning Approach Using Regularized Adjusted Plus-Minus (RAPM)", is strictly a product of my own personal work and all sources used have been duly stated in the bibliographic citations and references. Where I have used ideas, text and/or sources of other authors, they are clearly mentioned in the text with the appropriate citation and the relevant reference is included in the bibliographic references section with a full description.

The Declarant

Maksim Misevich

Abstract

Projecting which college basketball players will succeed in the NBA remains one of the most challenging tasks in professional sports. This study utilizes publicly available NCAA statistics to investigate whether machine learning models can predict players' rookie window impact, using Regularized Adjusted Plus-Minus (RAPM) over their first four NBA seasons as the proxy metric.

Using data on 546 drafted players and undrafted free agents from 2008 to 2021, the analysis applies extensive feature engineering including era-adjusted efficiency metrics, per-100 possessions standardization, team quality ratings and team strength of schedule. A temporal train (2008-2018) / test (2019-2021) split is employed to evaluate whether the model can generalize on new, unseen data. Two models are compared: a regularized linear model (ElasticNet) and a gradient-boosted tree ensemble (XGBoost), with hyperparameter tuning via cross-validation grouped by draft year.

Both models explain approximately 7-8% of the variance in the rookie window RAPM on the held-out test set. However, despite the modest R^2 values, both models demonstrated meaningful ranking ability, outperforming the aggregated pre-draft media consensus board (pooled Spearman ρ of 0.265 for XGBoost and 0.233 for ElasticNet vs 0.050 for consensus). Feature importance analysis shows that both models converge on defensive stops per 100 possessions, free-throw percentage, assists to turnover ratio and draft age as the strongest predictors of early career NBA impact.

The findings indicate that while college statistics provide a limited but real signal for NBA projection, interpretable models can still add value as a supplementary tool to traditional scouting by identifying undervalued or overvalued prospects relative to consensus expectations.

Keywords: NBA Draft, player projection, RAPM, machine learning, sports analytics, college basketball, regression.