# Bimal Kumal

917-821-6444 | bkumal@student.sjny.edu | linkedin.com/in/bimal-kumal | github.com/bimal1023

## SUMMARY

Computer Science & Data Science undergraduate with hands-on experience in Machine Learning, LLMs, and cloud-based MLOps. Skilled in building end-to-end ML and GenAI systems using Python, AWS, and LangChain.

## EDUCATION

**St. Joseph's University, New York** **Brooklyn, NY**
*BS in Computer Science & Data Science* Expected: May 2027
*GPA: 3.94*

**Relevant Coursework**: Data Structures & Algorithms, Advanced Programming, Operating Systems, Database Systems, Linear Algebra, Calculus I, II, Software Engineering, Probability & Statistics

## SKILLS

**Programming Languages:** Python, Java, SQL, R
**ML Framework**: Scikit-learn, XGBoost, TensorFlow, Keras, PyTorch, Hugging Face Transformer
**GenAI & LLMs**: LangChain, RAG, Fine-tuning (LoRA/QLoRA), Prompt Engineering
**Data Tools**: Pandas, NumPy, FAISS, Chroma
**MLOps & Cloud**: Docker, AWS (SageMaker), MLflow, Git

## PROJECTS

**LLM-Based Document Intelligence System (RAG+LangChain)** **Aug 2025**
*Tech Stack: Python, LangChain, RAG, FAISS, Chroma, OpenAI API, AWS*

- Architected a RAG pipeline using LangChain and OpenAI API, reducing manual document retrieval time by 80%.
- Optimized search relevance by 30% through the implementation of FAISS vector embeddings and semantic search over traditional keyword-based methods.

**Lung Cancer Prediction Using Machine Learning** **Apr 2025**
*Tech Stack: R, Random Forest, Statistical Analysis, Feature Engineering*

- Applied Kruskal-Wallis and Chi-square tests in R to analyze risk factors, streamlining feature selection and cutting model development time by 25%.
- Developed a Random Forest classification model in R to categorize lung cancer risk with high precision, utilizing patient clinical data and health indicators.
- Designed a real-time predictive dashboard using R Shiny, enabling healthcare providers to visualize severity probabilities and risk correlations.

**Calories Burned Prediction Using Machine Learning** **Jan 2025**
*Tech Stack: Python, Linear Regression, Data Preprocessing, Feature Engineering*

- Built a high-performance Linear Regression model achieving a 0.960 $R^2$ score and Mean Absolute Error (MAE) of 8.03.
- Executed end-to-end data engineering on a dataset of 15,000 records, including feature scaling and BMI feature engineering to boost model predictive power.

## EXPERIENCE

**St. Joseph's University, New York** **Sep 2023 - Dec 2023**
*Honors Peer Tutoring* *Brooklyn, New York*

- Provided one-on-one and group tutoring in pre-calculus and introductory computer programming, helping students improve their understanding and grades.
- Created tailored study plans to address individual student needs, resulting in a 20% improvement in average test scores.

## CERTIFICATION & AWARDS

- Esse Non Videri Scholar: Awarded 90% tuition scholarship for academic merit and leadership.
- Introducing Generative AI with AWS - Accenture
- Supervised Machine Learning (Andrew Ng): Stanford Online
- Dean's List (5 consecutive semesters): Awarded to the top 5% of students for academic excellence.

# Transcript

St. Joseph's University
NEW YORK



Issued to:

BIMAL KUMAL
,

St. Joseph's University
245 Clinton Ave.
Brooklyn, NY 11205-2688

Name: Kumal, Bimal Date: 02/12/2026
SSN: ###-##- ID: 0432166
Birth Date: 03/13/2004

UNDERGRADUATE TRANSCRIPT

| COURSE | Course Title | GRD | | CRD |
|---|---|---|---|---|
| | FALL 2023 (09/06/2023 to 12/19/2023) | | | |
| COM152 | COMPUTER PROGRAMMING | A | | 3.00 |
| COM200 | COMPUTER SCI: AN OVERVIEW | A | | 3.00 |
| ENG103H | WRITING EFFECTIVE COMM | A- | | 3.00 |
| MAT205 | CALCULUS/ANALYTIC GEOM I | A | | 4.00 |
| SJNY100H | FRSEM GREAT WESTRN IDEAS HONOR | B | | 3.00 |
| TERM CR: 16.00 | GPA CR: 16.00 | QP: 60.10 | GPA: | 3.756 |
| CUM CR: 16.00 | GPA CR: 16.00 | QP: 60.10 | GPA: | 3.756 |
| Dean's List | | | | |
| | SPRING 2024 (01/22/2024 to 05/11/2024) | | | |
| ECO120 | MACROECONOMICS | A | | 3.00 |
| HON100 | HONORS FAITH IN FILM | A | * | 1.00 |
| MAT203 | MATH FDTNS OF COMPUTER SCIENCE | A | | 4.00 |
| MAT206 | CALCULUS & ANALYTIC GEOM II | A | | 4.00 |
| MUS212 | LATIN AMERICAN MUSIC | A | | 3.00 |
| TERM CR: 15.00 | GPA CR: 15.00 | QP: 60.00 | GPA: | 4.000 |
| CUM CR: 31.00 | GPA CR: 31.00 | QP: 120.10 | GPA: | 3.874 |
| Dean's List | | | | |
| | FALL 2024 (09/04/2024 to 12/17/2024) | | | |
| COM145 | INTRODUCTION TO DATA SCIENCE | A | | 3.00 |
| COM210 | ALGORITHMS & DATA STRUCTURES | A | | 3.00 |
| MAT246 | PROBABILITY AND STATISTICS | A | | 3.00 |
| MAT356 | LINEAR ALGEBRA | A | | 3.00 |
| MUS100 | UNDERSTAND, ENJOY MUSIC | A | | 3.00 |
| TERM CR: 15.00 | GPA CR: 15.00 | QP: 60.00 | GPA: | 4.000 |
| CUM CR: 46.00 | GPA CR: 46.00 | QP: 180.10 | GPA: | 3.915 |
| Dean's List | | | | |
| | SPRING 2025 (01/22/2025 to 05/13/2025) | | | |
| COM245 | PRIN AND STAT METH OF DATA SCI | A | | 3.00 |
| COM249 | COMP ORGANIZ/ASSEMBLY LANG | A | | 3.00 |
| HON100 | THE ART & CRAFT OF PUPPETRY & | A | * | 1.00 |
| MAT241 | HISTORY OF MATHEMATICS | A | | 3.00 |
| SCI150 | INTRO. TO PHYSICAL SCIENCE | A | | 3.00 |
| SCI150L | LAB FOR SCI-150 | | | |
| SOC100 | INTRODUCTORY SOCIOLOGY | A | | 3.00 |
| TERM CR: 16.00 | GPA CR: 16.00 | QP: 64.00 | GPA: | 4.000 |
| CUM CR: 62.00 | GPA CR: 62.00 | QP: 244.10 | GPA: | 3.937 |
| Dean's List | | | | |

| COURSE | Course Title | GRD | | CRD |
|---|---|---|---|---|
| | FALL 2025 (09/03/2025 to 12/16/2025) | | | |
| COM250 | SCRIPTING LANGUAGES | A | | 3.00 |
| COM380 | DATABASE SYSTEMS | A | | 3.00 |
| HON100 | HONORS | A | R | 1.00 |
| PHI160 | INTRODUCTION TO ETHICS | A | | 3.00 |
| RS168 | WORLD RELIGIONS | A | | 3.00 |
| TERM CR: 13.00 | GPA CR: 13.00 | QP: 52.00 | GPA: | 4.000 |
| CUM CR: 75.00 | GPA CR: 75.00 | QP: 296.10 | GPA: | 3.948 |
| Dean's List | | | | |
| | SPRING 2026 (01/22/2026 to 05/13/2026) | | | |
| COM140 | COMPUTER APPLICATIONS I | CIP | | |
| COM310 | OPERATING SYSTEMS | CIP | | |
| COM370 | ADV.COMPUTER PROGRAMMING | CIP | | |
| MAT350 | DIFFERENTIAL EQUATIONS | CIP | | |
| POL102 | INTRO TO POLITICAL SCIENCE | CIP | | |
| TERM CR: 0.00 | GPA CR: 0.00 | QP: 0.00 | GPA: | 0.000 |
| CUM CR: 75.00 | GPA CR: 75.00 | QP: 296.10 | GPA: | 3.948 |

PROGRAM: Mathematics/Computer Science, B.S.
END OF TRANSCRIPT

Robert Pergolis
Registrar

Continued on next Column/Page

An official transcript is printed on copy safe paper, does not require a raised seal, and is valid only when it bears the signature of the appropriate college official. Copies issued to students will have "Issued to Student" printed on the transcript.

**Repeated Courses**

A student who receives an unsatisfactory grade in a course specifically required for the degree, for the major, or for a certificate program may request departmental approval to repeat the course Both the original grade and the second grade will appear on the transcript. Credit will be given only once for the course, and the most recent grade will be calculated in the index. (Symbols: asterik (*) forgiven grade, (R) most recent grade.)

**RELEASE OF TRANSCRIPTS**

In accordance with the Family Educational Rights and Privacy Act of 1974, the information contained on this transcript shall not be released to any other party unless written consent is obtained from the student.

**Personal Statement: Big Data Summer Immersion at Yale (BDSY)**

I am fascinated by how AI/ML can analyze medical images to identify early patterns in diseases like lung cancer, which claims millions of lives annually. As part of my data science project, I developed a Random Forest classification model in R to categorize lung cancer risk with high precision using patient clinical data and health indicators. The most challenging part of this project was that the dataset was not clinically verified by a hospital or cancer registry; therefore, it may not reflect the clinical relevance of actual patient outcomes. I am eager to apply my knowledge gained from my academic project to real-world projects that make a positive impact on society. Through this program, I look forward to gaining access to sufficient resources to analyze and extract meaningful insights.

Furthermore, I am excited to be involved in hands-on research guided by Yale faculty and graduate mentors and meet with the diverse community to collaborate on various projects to solve complex health problems from different perspectives. This opportunity will play a vital role in my goal of becoming a Machine Learning engineer or Data Scientist specializing in the healthcare industry. I have a strong foundation in data cleaning, wrangling, and AI, and I am ready to bring my technical skills and passion for medical data to the Yale community this summer. Outside of my coursework, I am currently pursuing certifications in AWS Machine Learning and Oracle Generative AI. These have taught me how to build end-to-end models, a skill which I hope to apply when working with Yale's massive health datasets.