



# PUNIT KUMAR

Aspiring Data Engineer



+91 86500 01977



punitkumar4871@gmail.com



[punitkumar4871](#)



<https://bento.me/punit-kumar>

## SKILLS

- **Programming Languages:**

Python, Java, JavaScript

- **Frontend Development:** HTML,

CSS

- **Backend & Databases:** MySQL,

MongoDB

- **Data Science & Analytics:**

Pandas, NumPy, Matplotlib,

Scikit-Learn, Power BI

- **Cloud Computing:** AWS, Azure

- **Version Control &**

**Collaboration:** Git, GitHub

## About Me

I am a student driven by a passion for solving complex challenges and creating impactful solutions. Keen to develop innovative systems that optimize efficiency and scalability. Recognized for a collaborative mindset, strong leadership abilities, and a commitment to personal and professional growth.

## EDUCATION

Aug 2023 - May 2027

Lovely Professional University, Jalandhar, Punjab

Bachelor of Technology (B.Tech)

- **CGPA: 8.69**

- Participated in coding event like Coding hunt, rapid fires and google developer groups events.
- Microsoft Learn Student Ambassador
- AWS Cloud Club Member

Apr 2021 - May 2023

St. Dominics Sr. Secondary School, Mathura, UP

Sr Secondary School

- **Percentage: 88.3%**

- Gold in Relay 200m rounds.
- Member of the social volunteering club in 2022
- Gold medal in basketball in inter school.

## EXTRACURRICULAR

- Open Source Contributor
- Volunteering for Environmental Initiatives
- Social Media Content Creator

# CERTIFICATIONS

---

- Microsoft certified: Azure Data Fundamentals (DP-900)
- IBM machine learning
- IBM Data Engineering
- Deep learning: Generative AI for everyone
- Geeks for Geeks: DSA in Python

# PROJECTS

---

## 1. FutureLMS

**Technology: Python, MySQL, Tkinter**

- Developed an intuitive learning management system to streamline course, grade, and attendance tracking for students and educators.
- Designed as a team project to enhance practical knowledge and deliver a functional solution for educational institutions.

## 2. Crop yield Prediction System

**Technology: Python, Azure, Scikit-learn, PySpark, Databricks, Gemini API, Github pipeline**

- Developed a dynamic data pipeline to ingest data from GitHub to Azure Data Lake Storage Gen2
- Preprocessed and cleaned large datasets using Databricks and PySpark
- Conducted data analysis using big data tools to identify key factors affecting crop yield
- Trained a machine learning model to predict crop yield based on processed data
- Implemented an end-to-end data engineering workflow integrating data ingestion, processing, analysis, and prediction

## 3. Real-Time Chat Application

**Technology: HTML/CSS, Flask, SocketIO, MongoDB**

- Developed a real-time chat app using Flask as the backend framework and Socket.IO for real-time bi-directional communication
- Integrated MongoDB to store and manage chat messages persistently
- Implemented user presence features including join and leave notifications
- Enabled multi-user interaction within shared chatrooms
- Designed and deployed a clean, responsive UI for cross-device compatibility

## 4. Healthcare Data Analysis in India

**Technology: Python, NumPy, Pandas, Seaborn, Matplotlib**

- Analyzed nationwide healthcare data to uncover trends in disease outbreaks, healthcare access, and population demographics.
- Utilized advanced data visualization to present actionable insights, supporting strategic improvements in public health.