# Sudhanshu Sharma

Email: ********@gmail.com
Phone: +91**********
LinkedIn: (Name) (https://www.linkedin.com/in/**********)
GitHub: (GITHUB Username) (https://github.com/********)

## Career Objective

Aspiring Data Scientist and GenAI Developer with a strong background in machine learning, large language models, and data analysis. Proficient in designing predictive models and developing intelligent systems to solve complex business problems. Passionate about leveraging data to drive impactful decisions and create innovative solutions.

## Key Skills

- Programming Languages: Python, SQL
- Data Science & Machine Learning:
  - Regression and Classification Models
  - Predictive Modeling
  - Statistical Analysis
  - Hypothesis Testing
  - Model Validation (Cross-Validation, ROC Curves)
  - Ensemble Learning (Random Forest, Gradient Boosting, XGBoost)
  - Feature Engineering
  - Model Deployment
- Deep Learning:
  - Artificial Neural Networks (ANN)
  - Convolutional Neural Networks (CNN)
  - Recurrent Neural Networks (RNN)
  - Transfer Learning
  - NLP Libraries (Spacy, NLTK, Gensim, Word2Vec, Transformers)
  - Generative Models (GANs, VAEs)
- Data Visualization:
  - Tableau
  - Power BI
  - Matplotlib
  - Seaborn
  - Plotly
- Web Development:
  - Flask

  - Streamlit
  - HTML/CSS
- Database Management:
  - SQL (Joins, Views, Procedures)
  - NoSQL (MongoDB)
  - MySQL
  - SQLite
- EDA (Exploratory Data Analysis):
  - Pandas
  - NumPy
  - Pandas Profiling
- Big Data Tools:
  - Hadoop
  - Spark
  - Hive
- Data Analysis Tools:
  - Excel
  - Jupyter Notebooks
  - Google Colab
- Version Control & Collaboration:
  - Git
  - GitHub
- Cloud Platforms:
  - AWS (S3, EC2, RDS)
  - Google Cloud Platform (BigQuery, AI Platform)
  - Microsoft Azure (Machine Learning Studio)
- GenAI & LLM Tools:
  - Langchain
  - HuggingFace
  - OpenAI
  - GPT-3, GPT-4
  - BERT, RoBERTa, GPT-Neo
- Other Skills:
  - EDA
  - Credit Risk Analysis
  - Project Management
  - Agile Methodologies
  - Data Storytelling

## Technical Skills

- Python:
  - Libraries: Pandas, NumPy, Scikit-Learn, Matplotlib, Seaborn, Plotly, TensorFlow, Keras, PyTorch, Dask, Pandas Profiling
  - Concepts: Data Manipulation, Data Cleaning, Feature Engineering, Model Training and Evaluation, Hyperparameter Tuning
- Machine Learning & AI:
  - Algorithms: Linear Regression, Logistic Regression, Decision Trees, Random Forest, SVM, KNN, Naive Bayes, Gradient Boosting, XGBoost, AdaBoost
  - Techniques: Cross-Validation, Grid Search, Random Search, Ensemble Learning, Feature Selection
  - Libraries: Scikit-Learn, XGBoost, LightGBM, CatBoost
- Deep Learning:
  - Frameworks: TensorFlow, Keras, PyTorch
  - Architectures: ANN, CNN, LSTM, GRU, GANs, VAEs
  - Libraries: Transformers (HuggingFace), TensorFlow Hub, PyTorch Hub , tensorflow , Data Augumentation(ImageDataGenerator), Keras(Pre trained models such as resnet, resnet 50, vgg16, vgg19,xception), Mediapipe(Object detection , landmark detection.)
- NLP:
  - Libraries: Spacy, NLTK, Gensim, Word2Vec, FastText, Transformers (BERT, GPT-3, GPT-4), HuggingFace
  - Techniques: Text Preprocessing, Tokenization, Lemmatization, Stemming, Named Entity Recognition (NER), Sentiment Analysis, Topic Modeling
- Data Visualization:
  - Tools: Tableau, Power BI, Matplotlib, Seaborn, Plotly, ggplot2 (R)
  - Techniques: Interactive Dashboards, Data Storytelling, Drill-Down Analysis
- EDA:
  - Libraries: Pandas, NumPy, Dask, Pandas Profiling
  - Techniques: Descriptive Statistics, Data Cleaning, Data Transformation, Correlation Analysis, Outlier Detection
- Database Management:
  - SQL: Joins, Views, Stored Procedures, Indexing, Optimization
  - NoSQL: MongoDB
  - RDBMS: PostgreSQL, MySQL, SQLite
- GenAI & LLM:
  - Tools: Langchain, HuggingFace, OpenAI, GPT-3, GPT-4, BERT, RoBERTa, GPT-Neo
  - Concepts: Fine-Tuning, Transfer Learning, API Integration, Agent-Based Models
- Web Development:
  - Frameworks: Flask, Streamlit, Django
  - Frontend: HTML/CSS, Bootstrap
- Big Data Tools:
  - Hadoop, Spark, Hive

- Cloud Platforms:
  - AWS: S3, EC2, RDS
  - Google Cloud Platform: BigQuery, AI Platform
  - Microsoft Azure: Machine Learning Studio
- Data Analysis Tools:
  - Excel, Jupyter Notebooks, Google Colab

## Projects

### Credit Risk Management

Objective: Developed a predictive model to assess credit risk and categorize customers based on risk levels.
Technologies: Python, Machine Learning, SQL, Tableau
Impact: Improved risk assessment accuracy and provided actionable insights for financial institutions.

### Diabetes Prediction

Objective: Early detection of diabetes using various classification algorithms.
Technologies: Python, Flask, Machine Learning
Impact: Enabled proactive healthcare management through early diagnosis.

### Parkinson's Disease Prediction

Objective: Built a classification model to predict Parkinson's disease from patient data.
Technologies: Python, Flask, Machine Learning
Impact: Facilitated early treatment by accurate and timely prediction of Parkinson's disease.

### GenAI Chatbot

Objective: Developed a chatbot using large language models to provide information and assistance based on user input.
Technologies: Python, Streamlit, Langchain, HuggingFace, OpenAI
Impact: Enhanced user experience by providing quick and accurate responses to queries.

## Academic Qualifications

- B.Tech in Computer Science and Engineering
  - Institute: Sardar Beant Singh State University
  - Current Status: Pursuing
- 12th Standard: 92%, PSEB, 2020
- 10th Standard: 79%, PSEB, 2018

## Personal Details

- Date of Birth: 24-07-2003
- Gender: Male
- Address: Begowal (Taragarh), Pathankot, Punjab 143534
- Languages: English, Hindi, Punjabi

## Received Certifications from Platforms

### (1) HackerRank

### (2) Codechief

### (3) LeetCode