Statement of Purpose – Vandit Goel

My fascination with the world of Data began in my first year of Bachelors. Determined to learn programming, I embarked on a self-guided journey into Java and ventured into online courses, supplementing my knowledge with a collection of projects. My first endeavour involved developing an Interactive Earthquake Map application to display real time earthquake data. As part of the final year major project, I built a Voice Controlled Robotic Nurse designed to minimize healthcare professionals' exposure to contagious diseases. The robot was capable of understanding commands in Hindi and English. This project introduced me to Cloud based development and Text-to-Speech technologies, igniting my passion for Machine Learning.

These experiences opened doors, eventually leading me to secure a position at a Fortune 500 company, named, United Health Group (UHG) as a Data Analyst. Here I have developed numerous algorithms to identify and process overpayments and fraudulent healthcare claims. One of my most notable projects involved developing a model that saved more than $6 million per month by detecting overpayments for implants and high-cost drugs. I have also worked on a data pipeline that parses 150,000 claims daily from multiple sources and stores the output in Cassandra. This data is then consumed by PySpark algorithms and machine learning models built using Spark MLib, TensorFlow, and scikit-learn. Additionally, I have contributed to organization-wide migration from on-premises applications to the cloud. In this role, I built an infrastructure using Apache Kafka, Cassandra, MySQL, Apache Spark, and machine learning models hosted on Microsoft Azure.

My work as a Data Analyst has exposed me to a wide range of technologies. However, it is my work with machine learning models that has truly ignited my passion for this field. I am now eager to take the next step in my educational journey and pursue a formal degree in data science. I am particularly interested in deepening my understanding of neural networks, natural language processing (NLP), and large language models (LLMs). I am excited to explore all dimensions of this dynamic field and contribute to its continued advancement.

I aspire to amass valuable industry experience, contributing my expertise as a Data Scientist in prominent tech companies such as Google, Microsoft, Databricks and Optum. My goal is to develop machine learning models that have a meaningful impact on people's lives. Given my experience in the healthcare sector, I believe ML has huge scope in improving the accuracy and efficiency of diagnosis and treatment, it can help to identify overpayments and frauds which ensures that healthcare resources are used efficiently and that patients are not paying too much for their care.

Moreover, I am deeply committed to addressing the educational disparity among underprivileged children. My experiences as a volunteer tutor for the National Service Scheme - Abha Program shed light on the significant educational accessibility gap. In the long run, I aim to leverage my skills and knowledge to create sustainable solutions that benefit these children and facilitate their access to quality education. One of the major concerns I realized while tutoring mathematics to a 6th standard girl was that even though she was good at mathematics she had trouble understanding the questions as the textbooks were in English and not her native language. Today we have

technologies like Live Text translation and Live text-to-speech in the palm of our hands. If classrooms have access to tools like these it will be much easier to translate and provide explanations of the concepts in the student's native language. ML will also help in identifying and addressing gaps like these much faster than I or her teachers did.

After exhaustive groundwork, I have realized that the MS in Data Science program offered by the University of Texas Arlington perfectly aligns with my area of interest. I am particularly interested in exploring subjects like Artificial Intelligence I (CSE 5360). Moreover, it would be an honour to study under the tutelage of Dr. Upendranatha Sharma Chakravarthy since I find their research work in Big Data analysis quite intriguing, especially the recent research involving Multi-layer network decoupling. I also look forward to being a part of the Data Science Club and Maverick Chess Club.

In summary, my journey, from an self-taught developer to a Data Analyst, has been marked by unyielding determination, self-motivation, and a relentless drive to learn and create. I now seek to elevate my career through formal education in Data Science and eventually utilize my expertise to drive positive change, particularly in the education sector for underprivileged children.