

311 Stasney Street, #1300  
College Station, TX 77840

October 1, 2020

Hiring Manager,  
Bayer's Pharmaceuticals division,  
Hanover, Whippany, NJ 07981

Dear Hiring Manager,

I am a PhD student in Electrical Engineering at Texas A&M University, College Station and I am writing to apply for the data science roles that opened up recently at your division, namely, **Bayer Pharmacovigilance Data Science Co-op Spring 2021** (Reference ID: 263541) and **Bayer Pharmacovigilance Data Science Intern Summer 2021** (Reference ID: 263524).

As part of my coursework in *Pattern Recognition* during my PhD, I have had the opportunity to build a strong foundation in advanced analytics and machine learning techniques including SVM, decision trees, random forests, boosting, neural networks and autoencoders. During my coursework in *Regression Analysis* and *Applied Statistics and Data Analysis*, I worked on dozens of modeling problems in R using multiple regression, outlier detection, parsimonious model selection, generalized linear regression, mixed models, ANOVA and time series analysis. Currently I hold a 3.82 cumulative GPA, and 3.88 GPA during my PhD.

I have approximately three years of programming experience in R and Python, with over nine years of total coding experience including MATLAB and C. When it comes to striking a balance between code readability and efficiency, I believe there is no alternative to experience. I always try and test multiple ways of coding a problem with focus on vectorization and corner cases. Please refer to the enclosed resumé which provides additional information about my coding experience in both of these languages as well as in shell scripting. I believe my projects and coursework in machine learning and bioinformatics has provided me a strong theoretical foundation to learn new modeling techniques.

At A&M, I have had the opportunity to work with inter-disciplinary teams, be it cancer genomics or crop sciences research. Even with engineering background, I surveyed quite a bit of biological literature to understand the research problems and propose meaningful hypotheses. Additionally, I constantly had to learn new quantitative domains from graphical models to information theory. I think the ability to select a right algorithmic approach to solve a problem is a valuable skill for interdisciplinary research. As an example, I proposed a gaussian mixture modeling algorithm that took advantage of prior information available in literature about drug target locations for cancer heterogeneity estimation. I believe being able to go out of my comfort zone to learn new concepts in other disciplines make me a uniquely qualified candidate for this position.

This co-op/internship provides me a unique opportunity to learn about state-of-the-art machine learning techniques used in the industry for oncology drug discovery and hone my data science skills in the process. More importantly, as someone who has lost a family member to cancer, I am excited to be part of a company developing innovative and broad solutions like Larotrectinib, which by blocking an upstream NTRK fusion event, has been clinically proven to be an effective treatment for many types of cancer in certain populations.

I am very interested in learning more about the referenced data science positions at Bayer Pharmaceuticals. At your convenience, I'd like to meet with you via zoom/skype and discuss more about my qualifications and the value I can bring to your organization. I can be reached via email/skype at **rkapr.analytics@gmail.com** or via phone at **979-398-2987**. I look forward to talking with you and appreciate your consideration.

Sincerely,

Rajan.

Kapoor.CoverLetter