TurboStrat

November 14, 2023

Opportunity in Data Science

We currently offer a short-term internship, where the structure of the internship closely resembles that of a university thesis, encompassing research activities, coding tasks, and more. By engaging in this experience, you will further develop your research and problem-solving abilities, strengthen your coding expertise, and enhance your overall understanding of data science and machine learning.

Generally, we expect our interns to benefit from the internship by:

- Developing and comparing various classic classification machine learning models, including logistic regression, SVM, decision tree, and random forest, utilizing a soccer dataset to accurately predict game outcomes.
- Implementing feature engineering techniques, such as ranking algorithms, to enhance model performance and effectively measured the resulting improvements.
- Conducting programmed simulation study to backtest and evaluate the effectiveness of different betting strategies based on the developed machine learning models.
- Performing in-depth analysis and research, resulting in the publication of valuable insights related to soccer game predictions and betting strategies.
- Building and deploying machine learning models for real-world use, ensuring seamless integration and practical application of the developed solution.

Minimum Qualification

- Currently enrolled in a university and pursuing a degree in mathematics, statistics, data science, computer science, or related fields.
- Experience with Python or R.
- Successful completion of introductory-level statistics and probability courses.
- A minimum GPA of 3.0 is required.

Please check out our **internship plan** to learn more about it. If you are interested, please email **team@turbostrat.com** with your resume/CV and your transcript (unofficial is okay).

TurboStrat LLC
20830 Stevens Creek Blvd 1132, Cupertino, CA 95014
www.turbostrat.com (509) 335-5555

Internship Program Plan
Data Science Track

Summer 2023

Plan and Contact Information:

Mentor:	TBA	**Duration:**	TBA
Email:	TBA	**Weekly Meeting:**	Every Friday 11 am EST

Description During the 8-week internship, you will work under the guidance of data science and machine learning expert. This will provide an opportunity to expand your knowledge and skills in the interdisciplinary field of data science and sports analytics & prediction.

This internship presents a good opportunity for you to gain hands-on experience and proficiency in essential techniques, tools, and technologies necessary for effective data analysis and the construction of predictive models. Specifically, you will be working with sports data, allowing you to leverage real-world examples to learn and apply these concepts.

The structure of the internship closely resembles that of a university thesis, encompassing research activities, coding tasks, and more. By engaging in this comprehensive experience, you will further develop your research and problem-solving abilities, strengthen your coding expertise, and enhance your overall understanding of data science and machine learning principles.

Minimum Qualification: Currently enrolled in a university majoring Math, Stats, DS, CS, or related fields, with a minimum 3.0 GPA .

Objectives:

- Domain Knowledge Building - (1) Understand what sports analytics does, (2) Understand how statistics and mathematics are applied in this field
- Familiar with Industry standard tech stacks for data analytics and visualization - (1) Learn to use Python and Jupyter Notebook, (2) Libraries of MatPlotLib, Seaborn, Pandas, Numpy, Scikit-learn, tensorflow, pytorch

Resources: You will be working under the mentorship of industry and/or academia expert, you will have the opportunity to learn how data science can be applied to sports analytics and predictions and practice new tech stacks.

Format: This internship offers a remote work arrangement, providing interns with flexibility in managing their time. However, it is expected that interns dedicate a total of [Number of] hours throughout the duration of the internship. To facilitate learning, a variety of resources such as videos and tech blog articles will be provided for interns to review and occasionally replicate.

To ensure effective progress tracking, interns are required to document their internship activities on a weekly basis, either in the form of a written work log or a comprehensive report. These updates will be periodically reviewed to evaluate the interns' progress and provide constructive feedback.

Final Project You shall choose at least one of the three tracks to work on,

1. data mining & visualization
2. building machine learning model
3. betting strategy optimization

Outcome Upon completion of the internship program, it is expected that you will compose a concise, blog-style article in one of the three tracks, showcasing content of qualified quality. The evaluation of your article will primarily be conducted by your mentor, who will assess various aspects including the following dimensions:

1. Quality of Work: Your mentor will evaluate the overall quality of your article, considering factors such as clarity, coherence, and the depth of your insights. It is important to demonstrate a strong understanding of the subject matter and present your ideas effectively.
2. Ability to Learn: Your capacity to learn and absorb new concepts and techniques will be taken into account. The mentor will assess your ability to grasp and apply the knowledge and skills gained during the internship, showcasing your growth and adaptability in the field.
3. Skills, including Tech Stacks and Problem Definition & Solving: Your proficiency in relevant technical tools, technologies, and problem-solving methodologies will be evaluated. This includes your ability to identify and define problems, as well as your skill in implementing appropriate solutions using the necessary tech stacks.

For interns who have produced at least one qualified piece of work, mentors will be delighted to provide recommendation letters upon request. The recommendation letters can serve as valuable endorsements of your achievements during the internship and can support your future endeavors in the field.

Schedule

Week No.	Topics	Sugested Readings
1	Overview - Sports analytics and prediction	
2	Python and its libraries for data science	
3	Machine learning in practice: classic ML models for classification problem and performance evaluation	
4	Deep dive: Data mining & visualization on soccer data	
5	Deep dive: Build a predictive ML model using soccer data	
6	Deep dive: Soccer betting, as a mathematical optimization, backtesting using simulation	
7	Deep dive: Feature engineering, ranking algorithms	
8	Final Project	