Investigating Correlations and Predictive Modeling for Crystal Structure in Lithium-ion Battery Cathode Materials

**Abstract**: This research proposal outlines a comprehensive approach for analyzing a dataset encompassing various chemical and physical properties of lithium-ion battery cathode materials. The study aims to uncover correlations among these properties and develop predictive models for crystal structure. Through a combination of statistical techniques, including linear and polynomial regression, as well as machine learning, the research seeks to enhance our understanding of lithium-ion battery performance.

**Introduction**: Lithium-ion battery cathode materials are critical for facilitating the de-intercalation of lithium ions within batteries. The unique characteristics of their crystal structures significantly influence the efficacy of lithium-ion de-intercalation, ultimately impacting the electrochemical performance of these batteries. This study aims to investigate possible correlations between various chemical and physical properties and to predict crystal structures using advanced analytical techniques.

**Data Analysis**: The initial phase of data analysis involved the generation of a heatmap, providing a holistic view of the dataset. This heatmap revealed potential correlations among different properties, serving as the foundation for further investigation. In this proposal, we highlight several noteworthy correlations:

a. Negative correlation between band gap and E above hull b. Negative correlation between band gap and formation energy c. Negative correlation between the possibility of forming a band structure and the number of sites d. Negative correlation between volume and the possibility of forming a band structure.

**Correlation Analysis**: To explore these correlations, the study will employ various statistical techniques, including 2D linear regression and polynomial regression. Specifically, we will focus on predicting the band gap from formation energy and E above hull values to assess the viability of approximating the band gap using these parameters. The assessment of goodness-of-fit will be carried out through parity plots and R² values.

**Crystal Structure Prediction**: The crystal structure is pivotal in understanding lithium-ion battery performance. To predict crystal structures, the study will leverage machine learning techniques, primarily the K-nearest neighbors (KNN) algorithm from the scikit-learn package. This approach will enable the classification of materials with known physical properties but unknown crystal structures. Furthermore, the research will quantify the accuracy of these predictions by evaluating testing and training errors.

**Significance and Implications**: This research holds significant implications for the lithium-ion battery industry, offering insights into the intricate relationships between various chemical and physical properties and crystal structures. The findings will contribute to a deeper understanding of lithium-ion battery cathode materials and may guide the development of materials with enhanced electrochemical performance.

**Conclusion**: This proposal outlines a systematic approach to analyze a dataset of lithium-ion battery cathode materials, aiming to uncover correlations and develop predictive models for crystal structures. The research methodology, which combines regression analyses and machine learning, promises to provide valuable insights into materials' electrochemical performance, potentially advancing lithium-ion battery technology.