

MaterialsZone - The Lean R&D Solution

Technical Exercise for Lead AI Engineer Position

Project Background

You are tasked with a project for a customer, a leading advanced composite materials company. The company aims to enhance the performance characteristics of its products by optimizing various material properties, with a particular focus on tensile strength due to its critical role in product durability and functionality.

Exercise Objective

As the Lead AI Engineer, you are tasked with analyzing the provided data to derive insights and make recommendations for experimental setups that could potentially optimize the tensile strength of new formulations. Your recommendations will guide the R&D team in their experimental design and testing phases, directly impacting the next generation of materials produced by the company.

Provided Data

You received two datasets in CSV format:

1. **Raw Materials Dataset (`raw_materials.csv`):** Contains details about each raw material used in the formulations, including properties such as melting point, density, and purity.
2. **Formulations Dataset (`formulations.csv`):** Includes data on different formulations that were tested for the product of the company. Each record contains:
 - Formulation identifier and name
 - Processing conditions (temperature, time, pressure)
 - Composition (percentages of different raw materials used)
 - Tensile strength (known for some formulations, unknown for others)

Note: the Tensile Strength is not given in the last 5 formulations. These will be used to test your model.

Tasks

1. **Data Exploration and Preprocessing:**
 - Conduct exploratory data analysis to understand the relationships between formulation components, processing conditions, and tensile strength.
2. **Feature Engineering:**
 - Develop new features that could be useful for predictive modeling.
 - Assess the importance of different features in relation to tensile strength.
3. **Predictive Modeling:**
 - Develop a model to predict the tensile strength of formulations using the features derived from the raw material and formulation data.
 - Use appropriate techniques, considering the nature of the data and the business objective.
 - Validate the model using suitable metrics to assess performance and generalizability.
 - Provide a prediction for the tensile strength for the 5 last formulations in

the `formulations.csv` file for which the tensile strength is not given.

4. **Insight Generation and Recommendations:**

- Provide insights into how different raw materials and processing parameters affect tensile strength.
- Based on the model's findings, recommend specific experiments that the R&D team should conduct to optimize tensile strength using the existing raw materials. These recommendations might include adjustments to raw material compositions, processing temperatures, times, pressures, etc.

5. **Reporting:**

- Prepare a report (e.g. in Jupyter Notebook) documenting your analysis, methodology, results, and recommendations.
- Include visualizations that clearly communicate your findings and support your recommendations.

6. **Productization Proposal:**

- Suggest how this capability could be productized to interact with researchers to guide their work, accelerate experimentation, and enhance material development through user-friendly interfaces and real-time data analysis capabilities.
- No implementation is expected as part of this task. A theoretical design using a diagram, wireframe, mockup or any other means you find suitable is sufficient.
- Consider aspects such as API integration, user interface, data input/output, and real-time interaction.

7. **Presentation:**

- You will be asked to present your approach, code, key insights, and recommendations over a 1-hour call. The presentation should effectively communicate technical content to both technical and non-technical audiences.

Evaluation Criteria

- **Analytical Skills:** Effectiveness in handling data and deriving insights.
- **Accuracy of Predictive Models:** Ability to predict tensile strength reliably.
- **Innovativeness of Solutions:** Creativity in feature engineering and experimental design recommendations.
- **Coding Skills:** Proficiency in programming, ability to write clean, efficient, and well-documented code.
- **Strategic Thinking:** Ability to connect technical solutions with broader product and business goals, demonstrating how predictive modeling can be productized to enhance user interaction and accelerate R&D.
- **Clarity and Persuasiveness of Presentation:** Ability to convey complex ideas clearly and persuade stakeholders of the value of your recommendations.

This exercise is designed to simulate a real-world scenario where data-driven insights can lead to significant improvements in product development.

Good Luck!