Carnegie Mellon University

Experience, Skill Sets, and Technical Presentation

Yunhao Xu

Email: yunhaox@andrew.cmu.edu

Education and Working Experience

Master's Degree: Carnegie Mellon University, Materials Science and Engineering, 2026
- Advisor: Michael Bockstaller

Bachelor's Degree: Shanghai Institute of Technology, Materials Science and Engineering, 2024
- Advisor: Yingqiang Zhang

Internship:
- Ocean Reviver, 2025/06 - 2025/08 [Irvine, CA]
- Shanghai Chest Medical Devices Co., 2024/07 - 2024/08 [Shanghai, China]

Research Fields

HANDS-ON EXPERIMENTAL

Synthesis: Suspension Polymerization, ATRP, Thin Film Coating…

Characterization: TEM, SEM, DSC, DMA…

Testing: Mechanical, Chemical…

AI/ML

Polymer Property Prediction

ML Models: XGBoost, Random Forest, KNN…

Optimization: Hyperparameter Tuning, Cross Validation

SIMULATION & MODELING

2D & 3D Modeling: Auto CAD, SolidWorks

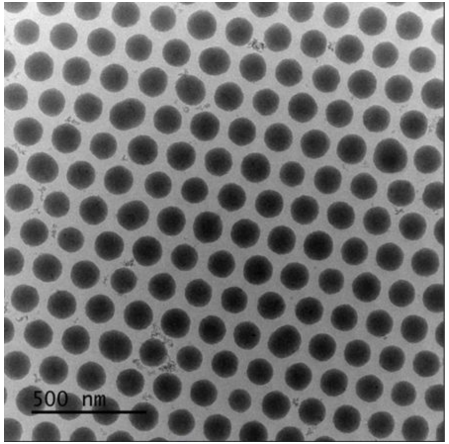
FEA Simulation: ANSYS

Overview of Research Projects

- Project 1 - Thin Film Preparation of Self Healing Hybrid Materials
- Project 2 - ML-Based Prediction of Mechanical Properties in Polymer-Grafted Nanoparticles
- Project 3 - Preparation and Performance Testing of Water-Based Acrylate Pressure-Sensitive Adhesives

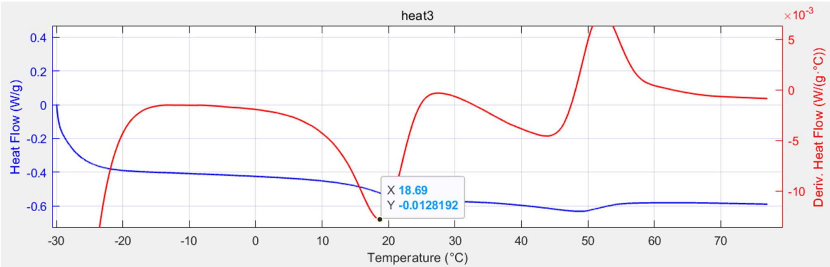
Skill Set: Advanced Characterization Techniques

Microstructural Analysis



- TEM (Self Healing Hybrids)
- Morphology

Performance Testing



- DSC, DMA, etc.
- Data Analysis: Matlab, Python, etc.

Skill Set: Computational Materials Science & AI

Data Acquisition (Experimental) → **ML Pipeline (Python)** → **Model & Sim (CNN)** → **Predictive Results (Heatmap)**

Data Acquisition (Experimental)

data
- pmma_data.csv
- ps_data.csv

	N	S	E
1	0.482	0.589	0.6784
2	0.763	0.629	0.8165

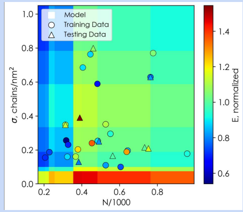
ML Pipeline (Python)

- generate_plots.py
- generate_tables.py
- train_models.py

Model & Sim (CNN)

saved_models
- DT.pkl
- KNN.pkl
- MLP.pkl
- OLS.pkl
- PR.pkl
- RF.pkl
- SVM.pkl
- XGB.pkl

Predictive Results (Heatmap)



Skill Sets and Lessons from Industry Research

Skill Sets (Medical Device R&D):
- CAD Design
- Standardized Testing
- QA & Root Cause Analysis

Skill Sets (Clean Energy & Tech):
- Direct Lithium Extraction (DLE)
- Quantum Computing Sim
- Process Optimization

Lessons:

Cross-Functional Collaboration

Bridging R&D with QA

Translating tech data into actionable insights

Adaptability

Transitioning industries

Rapidly learning new tech (Quantum, DLE)

Data Driven Impact

Real-world data is messy

Automation saves time

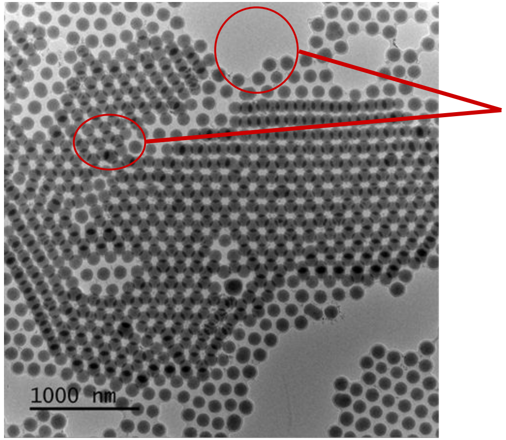
Industrial Mindset

Result-Oriented Timeline & Deliverables

Economic & Cost Consciousness

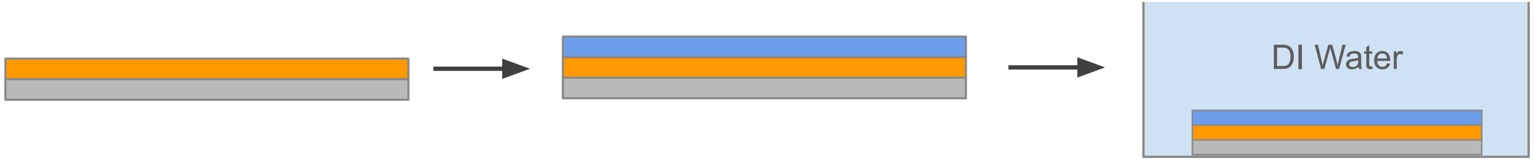
Background on the 1st project

Motivation:



Holes & Overlaps

Drop Cast

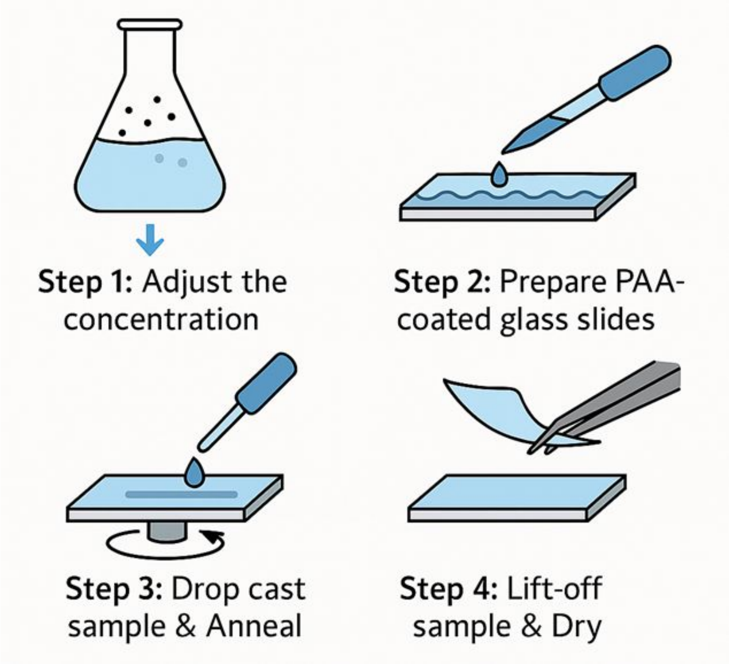


Substrate + PAA Layer

Spin Coating
oSiO2-PMMA

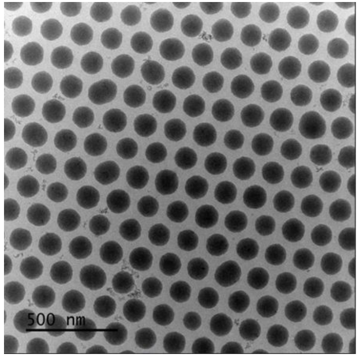
Lift Off (Sacrifice PAA)

Drop casting & Spin Coating



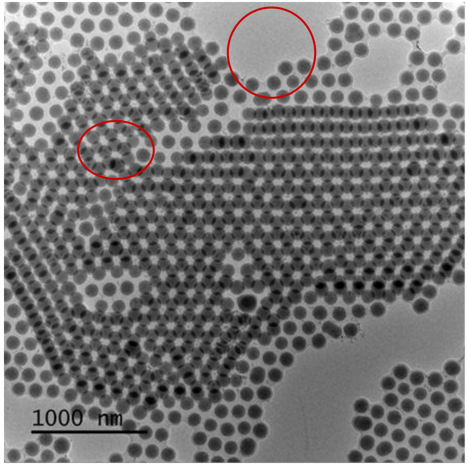
Key Variables:
1. Concentration
2. Spin Coating Speed & Time
3. Amount of Droplets

TEM Imaging:

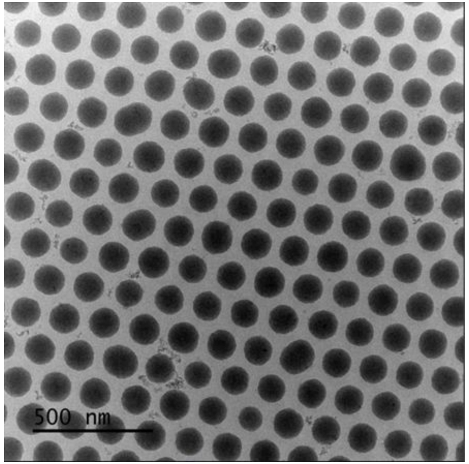


Impacts and Applications of 1st Project

Improving thin film quality:
Significantly decrease overlap and holes
Quality good enough for further research



Drop Cast

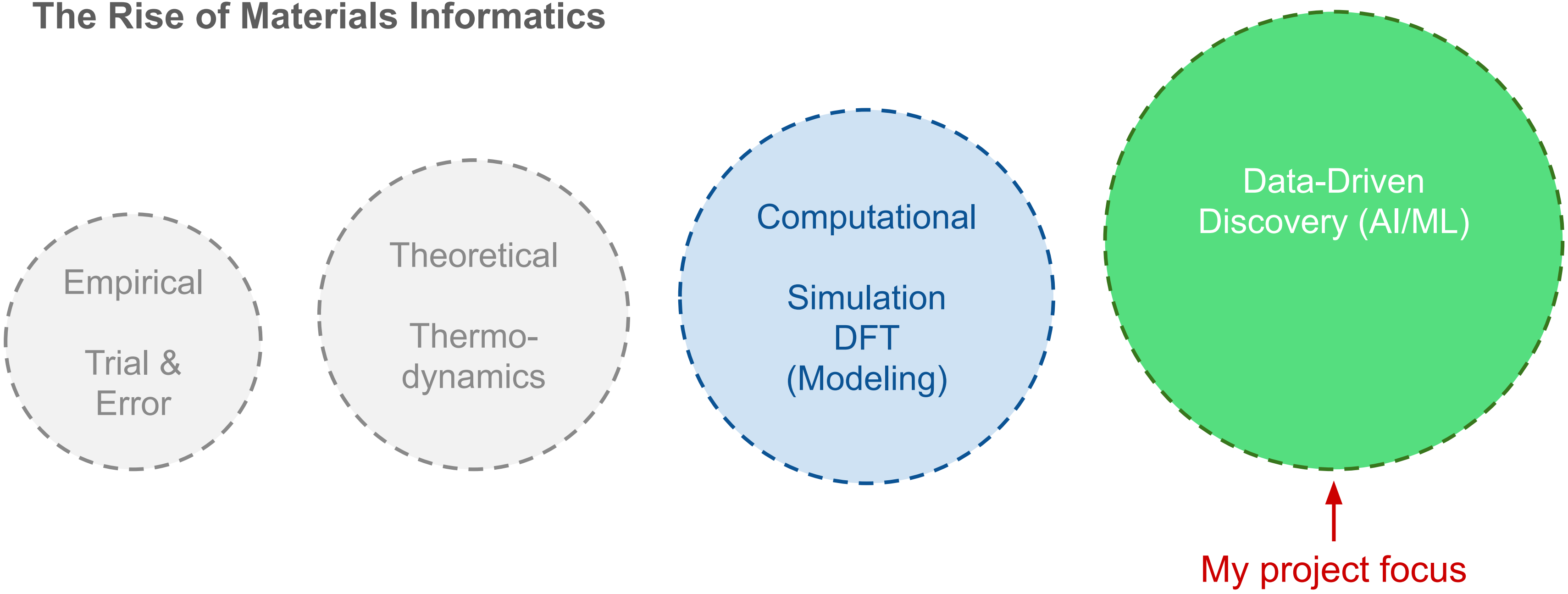


Spin Cast

Publication: Manuscript Completed & Under Review

Background on the 2nd project

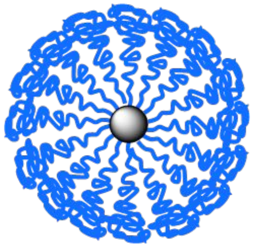
The Rise of Materials Informatics



Background on the 2nd project

ML-Based Prediction of Mechanical Properties in Polymer-Grafted Nanoparticles

Material System:
Polymer-Grafted Nanoparticle



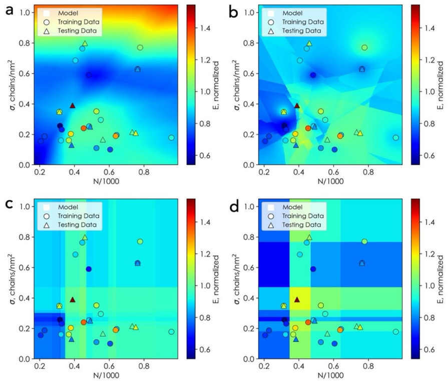
Key Parameters:

- N: Degree of Polymerization
- σ: Grafting Density

The Challenge

- Non-linear Relationship
- Experiments are costly

ML Approach



Target Output:

- Young's Modulus (E)

Goal: Predict material stiffness (Young's modulus)

Inputs:

- N - Degree of Polymerization (chain length, measured in units of 1000)
- σ - Grafting Density (chains/nm^2)

Output:

- E - Normalized Young's Modulus

Datasets:

PMMA Dataset - 42 data points

- Used for training (80%) and testing (20%)
- Split using a fixed random seed (49163) for reproducibility
- Features range: $N \in [0.206, 0.957]$, $\sigma \in [0.100, 0.799]$

PS Dataset - 34 data points

- Used exclusively for testing/validation
- Provides cross-polymer generalization assessment
- Young's modulus values normalized by 3.806 GPa

Machine Learning Models:

- OLS - Ordinary Least Squares (Linear Regression)
- PR - Polynomial Regression with Ridge regularization
- DT - Decision Tree Regression
- KNN - K-Nearest Neighbors Regression
- SVM - Support Vector Regression (RBF kernel)
- RF - Random Forest Regression
- MLP - Multi-Layer Perceptron (Neural Network)
- XGB - XGBoost Regression

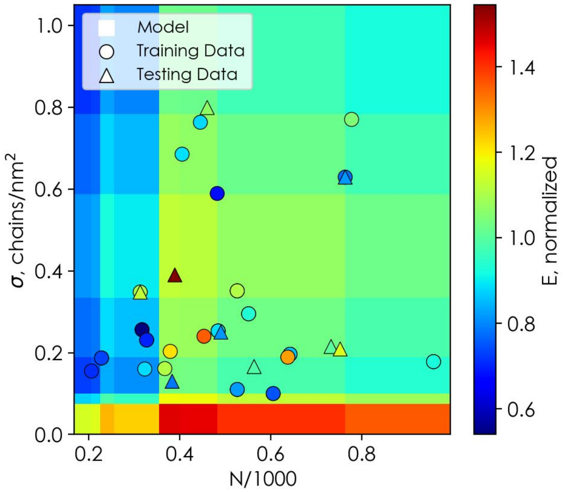
Workflow:

1. Model Training – Load PMMA data, perform LOO-CV grid search to minimize MAPE, train final models with best hyperparameters, save to saved_models/.
2. Model Evaluation – Load trained models, compute RMSE, MAPE, R^2, and export tables to tables/.
3. Visualization – Generate PMMA & PS heatmaps (● train / ▲ test) and PS error boxplots, save to figures/.

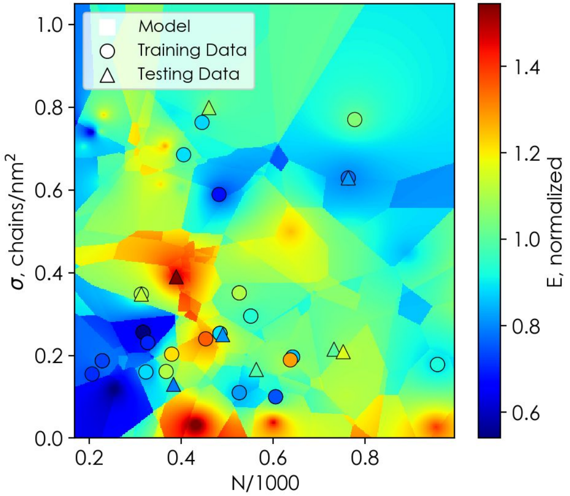
Results

	Model	Train MAPE (%)	Train RMSE	Train R²	Test MAPE (%)	Test RMSE	Test R²	PS MAPE (%)	PS RMSE	PS R²
1	DT	20.81	0.200	-0.122	20.34	0.250	-0.267	20.52	0.238	0.279
2	KNN	8.20	0.134	0.497	8.10	0.135	0.634	5.22	0.115	0.831
3	MLP	22.32	0.208	-0.213	17.12	0.236	-0.130	22.78	0.258	0.152
4	OLS	22.54	0.210	-0.233	16.69	0.234	-0.108	24.21	0.269	0.080
5	PR	22.47	0.209	-0.221	16.85	0.235	-0.119	22.80	0.256	0.163
6	RF	18.28	0.178	0.109	17.78	0.230	-0.064	17.69	0.187	0.553
7	SVM	22.73	0.211	-0.241	15.60	0.224	-0.016	25.35	0.277	0.021
8	XGB	18.48	0.176	0.134	16.06	0.206	0.140	16.81	0.184	0.570

XGB:



KNN:



Impacts and Applications of 2nd project

Efficiency & Cost

Traditional: ~2 weeks

< 1sec ML Prediction

Academic Deliverable

- ✔ Draft Completed
- Active — Scope Expansion
- Manuscript Submission

Background on the 3rd project

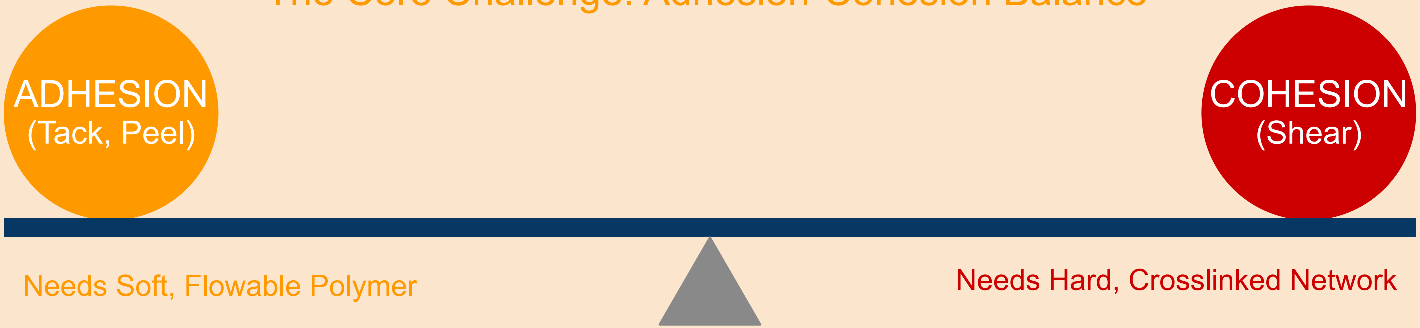
Why Waterborne?

- Eco-friendly (Low VOC)
- Safe (Non-flammable)
- Low Cost (Water solvent)

Key Components

- Soft Monomers (BA, 2EHA)
- Hard Monomers (EHMA, MMA)
- Functional Monomers (AA)
- Dispersant (PVA-1788, PVA-1799)
- Crosslinker (DAAM+ADH)
- Initiator (BPO)

The Core Challenge: Adhesion-Cohesion Balance

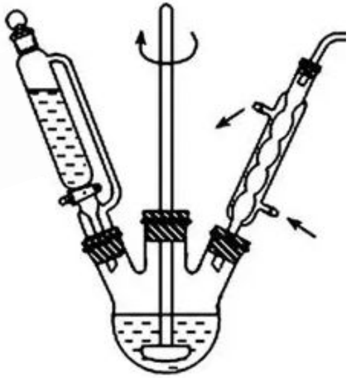


Synthesis & Experimental Design

Method: Suspension Polymerization (Core-Shell Structure)

Formulation Variables:

Soft/Hard Monomer Ratio	1.3, 1.7, 2.1, 2.6
Func. Monomer (AA)	0.3g, 0.6g, 0.9g
Initiator (BPO)	0.3g, 0.45g



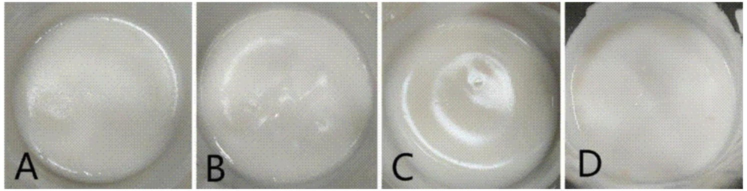
Process:

1. Setup: 500mL four-neck flask, reflux condenser, separatory funnel
2. Pre-mix: PVA / DI Water (20mins RT → 20mins 80°C)
3. Core: Add monomer → 1h reaction
4. Shell: Add monomer → 2–3h reaction
5. Output: Cool & Collect

Multi-Dimensional Characterization

Physical & Processing Properties

Appearance: Emulsion Stability



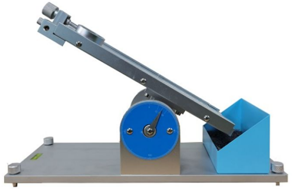
Viscosity: Rotational Viscometer

Solid Content: Thermal Oven Method

$$X = \frac{m_1}{m} \times 100$$

Adhesive Performance

Tack: Rolling Ball Method



Shear Adhesion: Cohesive Strength

180° Peel Strength: Tensile Tester

Results

AA (Acrylic Acid) & Crosslinking Density:

- Higher AA dosage → Increased Crosslinking Density.
- Impact: Excessive crosslinking restricts chain mobility, leading to a decrease in Peel Strength.

Core-Shell Structure & Tg:

- Hard-shell design results in a higher Tg, which generally limits tack.
- Optimal Point: Best Tack was achieved at an AA dosage of 0.6g.

Hard Monomer Ratio:

- Decreasing hard monomer ratio → Peel Strength↑ and Tack↑.
- Diminishing Returns: Beyond a certain threshold, the growth rate of both Tack and Peel strength slows down.

Thank you!
Questions?

yunhaox@andrew.cmu.edu