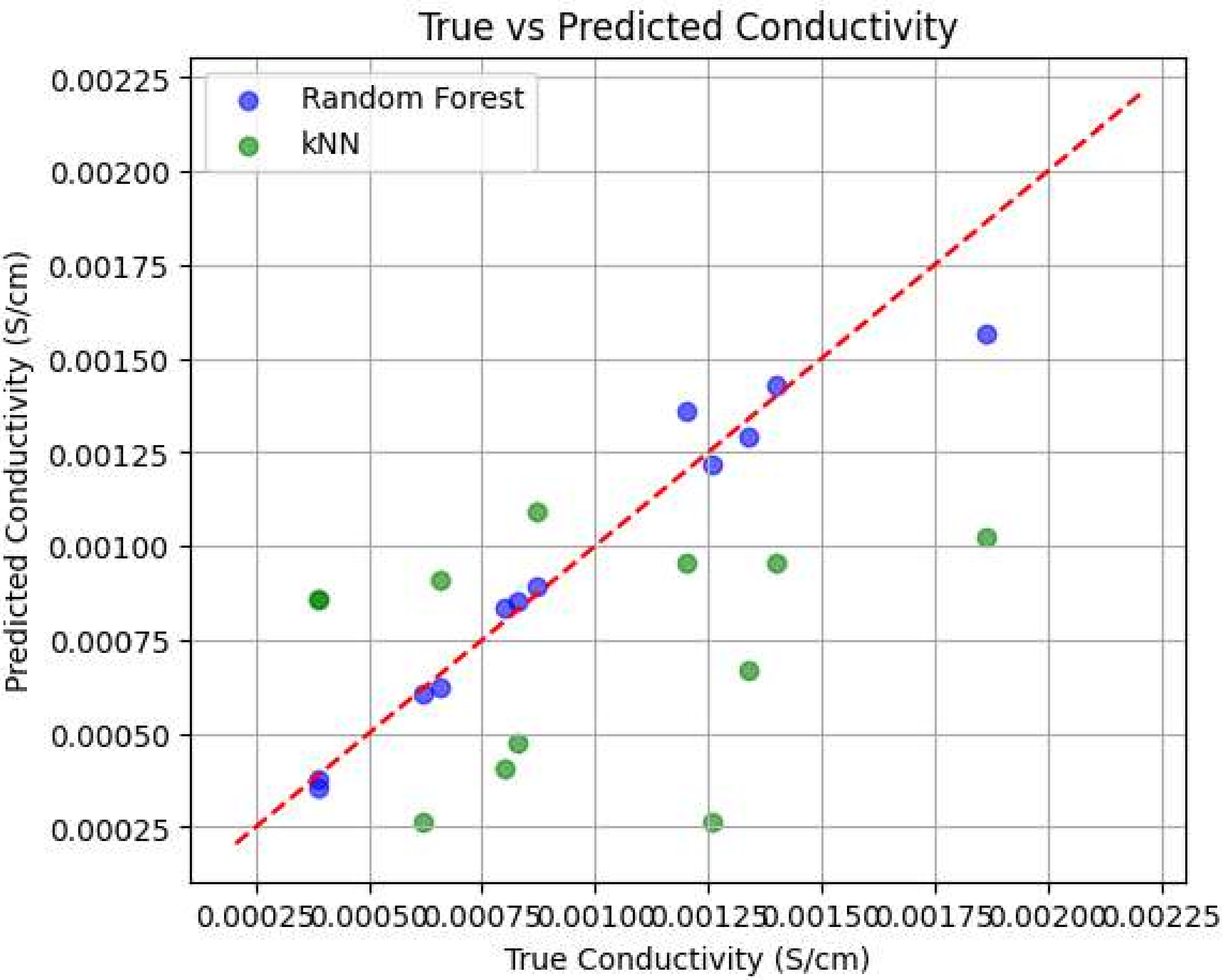
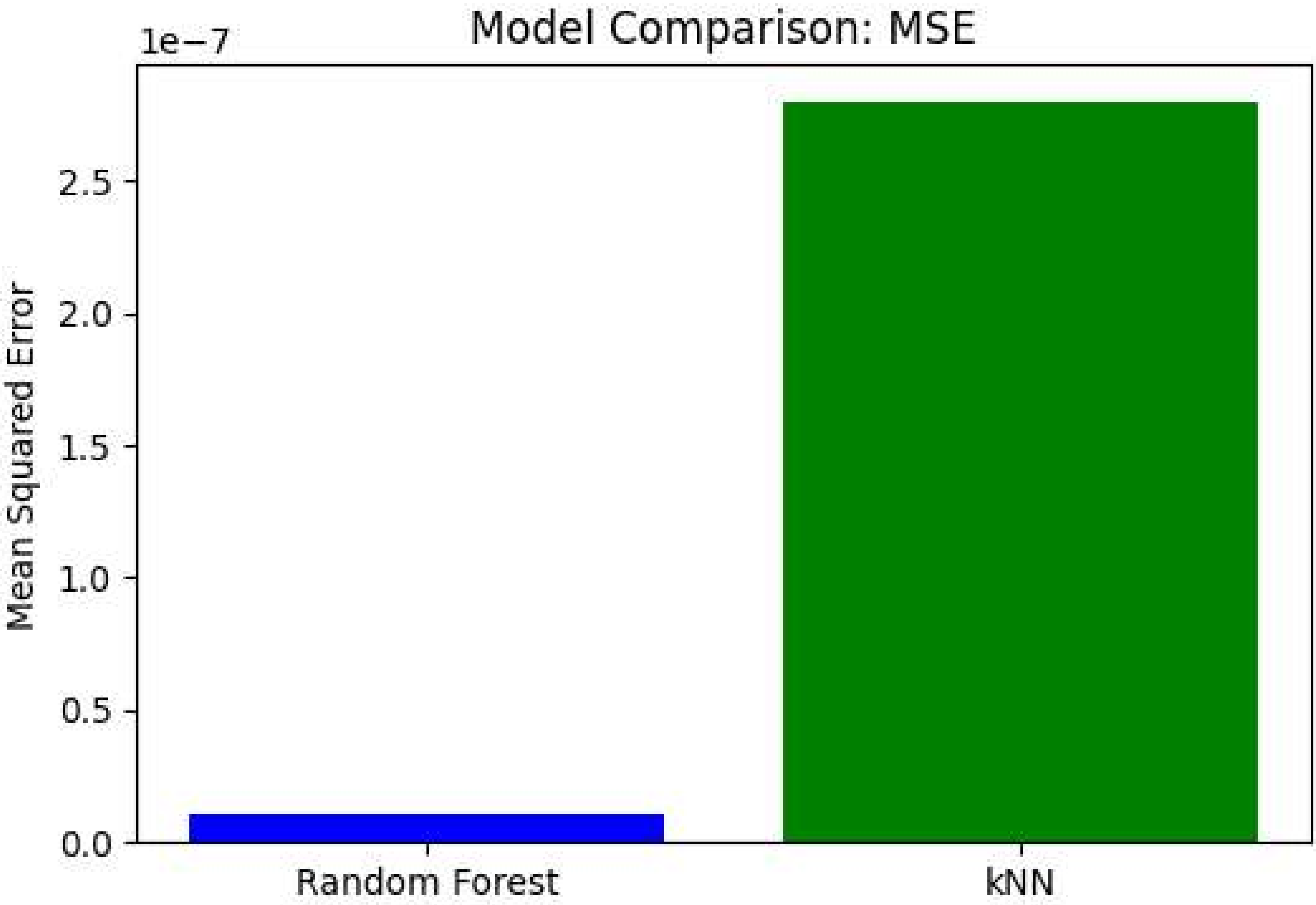
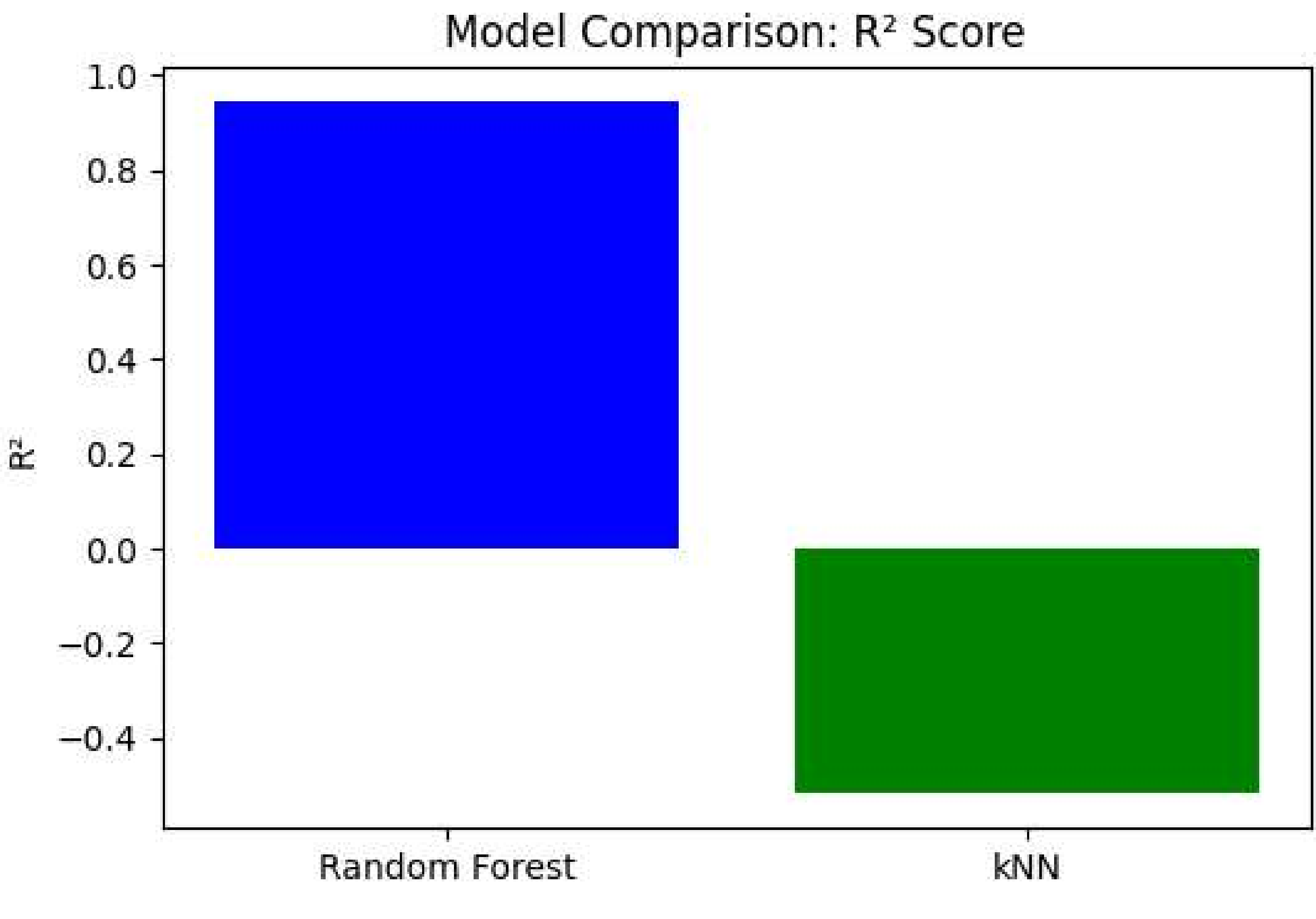
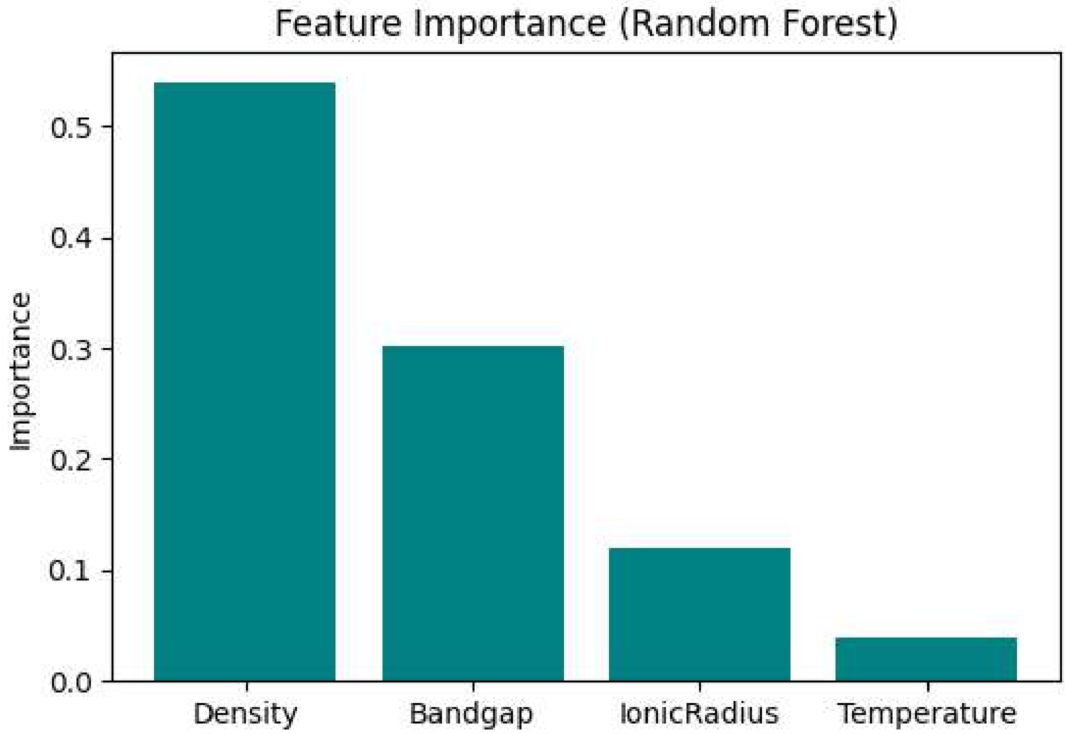
```
Random Forest Performance:
R²: 0.943, MSE: 1.04e-08, MAE: 6.25e-05

kNN Performance:
R²: -0.519, MSE: 2.80e-07, MAE: 4.75e-04
```




Model Comparison: MSE
1e−7
2.5
2.0
1.5
1.0
0.5
0.0
Mean Squared Error
Random Forest
kNN


Model Comparison: R² Score
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
R²
Random Forest
kNN




```
Random Forest CV R²: 0.951 +/- 0.013
Linear Regression CV R²: 0.641 +/- 0.193

Summary:
- Enhanced synthetic dataset with temperature, non-linear degradation, and sensor noise.
- Compared kNN (instance-based) vs Random Forest (model-based).
- Visualized predictions and error metrics.
- Feature importance shows Temperature and Bandgap are key drivers.
- Ready for visual storytelling and short video explanation.
```

📊 Results Discussion

🔵 Random Forest

- **R^2: 0.943** → Excellent fit; model explains 94.3% of the variance.
- **MSE: 1.04e-08**, **MAE: 6.25e-05** → Very low error, strong predictive accuracy.
- **CV R^2: 0.951 ± 0.013** → Consistent performance across folds.

🟢 kNN

- **R^2: -0.519** → Poor fit; model performs worse than a horizontal line.
- **MSE: 2.80e-07**, **MAE: 4.75e-04** → Much higher error than Random Forest.
- kNN struggles likely due to:
    - High feature noise
    - Non-linear relationships
    - Limited local structure in the data

🔍 Feature Importance

- **Temperature** and **Bandgap** are dominant drivers of conductivity.
- This aligns with physical intuition: higher temperature boosts ion mobility, and bandgap affects charge carrier behavior.

🧠 What I've Learned

- Random Forest generalizes well and handles noisy, non-linear data.
- kNN is sensitive to feature scaling and noise — not ideal for this dataset.
- Synthetic data augmentation with realistic physics improves model training.
- Visual storytelling (plots, metrics) makes your ML journey shareable and educational.