

# UV Index to SPF Recommendation – Final Report

## Project Overview

This project aims to build a predictive model that recommends suitable SPF levels based on the UV index, skin type, and exposure type. The goal is to aid skincare consumers in choosing the right sunscreen protection by understanding UV radiation risks.

---

## Data Description

The dataset includes the following features:

- **UV\_Index:** Numeric UV intensity on a scale from 0 to 11+
  - **Skin\_Type:** Categorical (Type I to Type VI based on Fitzpatrick scale)
  - **Exposure\_Type:** Categorical (e.g., Short, Moderate, Long exposure)
  - **Recommended\_SPF:** Categorical (e.g., SPF 15, SPF 30, SPF 50+)
- 

## Data Preprocessing

- **Encoding:**
  - **Skin\_Type** and **Exposure\_Type** were label encoded.
  - **Recommended\_SPF** values (e.g., "SPF 50+") were mapped to numeric SPF values like 50.
- **Null Handling:**
  - Nulls were found in the **Recommended\_SPF** column and removed.
- **Final Shape:**

- Training data shape: (400, 3)
  - Features used: UV\_Index, Skin\_Type\_Encoded, Exposure\_Type\_Encoded
  - Target: SPF\_Value (numeric)
- 

## Exploratory Data Analysis (EDA)

### Key Observations:

- **Correlation Matrix:**
    - Moderate correlation between UV\_Index and SPF values.
    - Lower correlation from Skin\_Type and Exposure\_Type, but still informative.
  - **Top Trends:**
    - **Higher UV Index** → Higher SPF recommended.
    - **Skin Type I/II** (fair skin) → Higher SPF.
    - **Longer exposure** → Higher SPF levels suggested.
- 

## Model Building

- **Model Used:** RandomForestRegressor
  - **Train-Test Split:** 80% - 20%
- 



## Evaluation Metrics

- **MAE** (Mean Absolute Error): Low → indicates accurate predictions
- **RMSE** (Root Mean Squared Error): Low → confirms model is not heavily penalizing outliers
- **R<sup>2</sup> Score**: High (close to 1) → indicates good model fit

MAE: 1.19

MSE: 2.54

RMSE: 1.59

R2 Score: 0.94

---

## Prediction Visualization

A comparison between predicted and actual SPF values:

Interpretation: The predicted SPF values closely follow the actual ones, suggesting a reliable model.

---

## Insights

1. **Small UV Index Changes** (e.g., 5 to 7) can cause SPF shifts from 15 to 30, especially for light skin types.
2. **Long Exposure** conditions always demand higher SPF.
3. **Skin Type VI** (dark skin) still requires SPF 15+ when UV Index is high.

---

## Final Takeaways

- We successfully built a regression model that predicts SPF level with over 94% accuracy.
- The model generalizes well on unseen data.
- Future improvements may include:

- Incorporating humidity and altitude data
- Adding real-time UV index inputs via API
- Building a user-friendly Streamlit or web dashboard