**Propensity Score Matching (PSM): Step-by-Step Documentation with Math and Explanation**

### **Objective:**

Estimate the causal effect of a binary treatment (e.g., using VM vs. Desktop) on a continuous outcome (e.g., application slowness in seconds), adjusting for covariates using Propensity Score Matching (PSM).

### **Step 1: Define Treatment, Outcome, and Covariates**

* **Treatment Variable (D):**
  + D\_i = 1 if unit i used a VM
  + D\_i = 0 if unit i used a Desktop
* **Outcome Variable (Y):**
  + Y\_i = observed application slowness (in seconds)
* **Covariates (X):**
  + X\_i = [Region, CPU type]
  + Encoded as dummy variables:
    - User\_region\_US-East
    - User\_region\_US-West
    - CPU\_type\_Intel\_i7
    - CPU\_type\_Ryzen\_5

### **Step 2: Estimate Propensity Score (e(X))**

The **propensity score** is the probability of receiving the treatment given covariates:

Estimated using **logistic regression**:

In code:

ps\_model = LogisticRegression(solver="liblinear")  
ps\_model.fit(X\_psm, y\_psm)  
df\_encoded["propensity\_score"] = ps\_model.predict\_proba(X\_psm)[:, 1]

✅ **Interpretation**: The higher the score, the more likely a unit is to be in the treated group (VM), given its covariates.

### **Step 3: Nearest Neighbor Matching on Propensity Score**

Match each treated unit (D=1) to the **closest control unit (D=0)** based on the propensity score:

In code:

nn = NearestNeighbors(n\_neighbors=1)  
nn.fit(control[["propensity\_score"]])  
\_, indices = nn.kneighbors(treated[["propensity\_score"]])  
matched\_control = control.iloc[indices.flatten()]  
matched\_control.index = treated.index

Each treated unit gets one matched control unit.

### **Step 4: Estimate ATT (Average Treatment Effect on the Treated)**

The ATT is the **average difference in outcome** between treated and matched controls:

In code:

att\_psm = (treated["slowness"] - matched\_control["slowness"]).mean()

✅ **Interpretation**: This tells us how much **slower** (or faster) the application is on VMs **compared to matched Desktops**.

### **Summary of Terms**

| Term | Meaning |
| --- | --- |
| X\_i | Covariates for user i (e.g., region, CPU) |
| D\_i | Treatment indicator (1 = VM, 0 = Desktop) |
| Y\_i | Observed outcome (slowness) |
| e(X\_i) | Propensity score = probability of being treated given X |
| ATT | Average treatment effect for treated units only |

### ✅ Final Output

If ATT = 0.415 seconds:

Users on VMs experience **0.415 seconds more slowness** than if they had used Desktops, after adjusting for region and CPU via PSM.

Let me know if you’d like the same treatment for **CEM**, **IPW**, or **T-/S-/X-learners**.