**📖 Theoretische Erklärung – ML-Ansatz**

**1. Ziel der Vorhersage**

Wir wollen für jedes Land (**Country\_ID**) die jährlichen **Total losses** (wirtschaftliche Verluste durch Naturkatastrophen) für die nächsten 5–10 Jahre schätzen.

**2. Welche Daten werden verwendet**

* **Zielvariable:**  
  Total losses aus deiner Datei Merged\_Energy\_Losses.csv
* **Prädiktoren (Features):**
  + **Zeitmerkmale**
    - Year (Jahr selbst)
    - year\_centered (Jahr relativ zum ersten Jahr → vermeidet zu große absolute Zahlen)
    - year\_sq (quadratischer Trend, falls es eine Beschleunigung/Verlangsamung gibt)
  + **Lagged Features (Verzögerungen):**
    - loss\_lag1, loss\_lag2, loss\_lag3 (Verluste der letzten 1–3 Jahre → Zeitreihenmodell-Charakter)
  + **Rollende Durchschnitte:**
    - loss\_roll3\_mean, loss\_roll5\_mean → glätten Ausreißer, geben mittelfristigen Trend
  + **Exogene Variablen:**
    - Alle Spalten, die mit PEC (Primärenergieverbrauch) oder FEC (Endenergieverbrauch) beginnen.  
      → erlaubt, Energieverbrauch als erklärende Variable einzubeziehen

Damit haben die Modelle sowohl:

* **historische Muster** (Lag-Features),
* **zeitliche Trends** (Jahr, Jahr²),
* **exogene Einflüsse** (Energieverbrauch).

**3. Warum gerade diese Modelle?**

**a) Ridge-Regression**

* Eine lineare Regression mit **L2-Regularisierung**
* Vorteil: schnell, stabil, liefert interpretierbare Koeffizienten
* Nützlich, wenn Daten eher lineare Trends aufweisen
* Weniger anfällig für Overfitting bei vielen korrelierten Features

**b) Random Forest Regressor**

* Ensemble aus vielen Entscheidungsbäumen
* Vorteil: erkennt nichtlineare Muster und Interaktionen
* Robust gegen Ausreißer und fehlende Normalverteilung
* Liefert solide Prognosen auch bei komplexeren Zusammenhängen

**4. Modellauswahl**

Für jedes Land wird **beide Modelle** trainiert und mit **TimeSeriesSplit Cross-Validation** verglichen.  
Das Modell mit dem **niedrigsten MAE (Mean Absolute Error)** wird gespeichert.  
So wird für Länder mit linearen Trends meist Ridge genommen, bei komplexeren Mustern Random Forest.

**5. Vorhersage-Logik**

* Für das jeweils letzte bekannte Jahr werden die Features berechnet
* Dann werden für jedes Jahr Schritt für Schritt neue Features aktualisiert  
  (z. B. loss\_lag1 = Vorhersage des Vorjahres)
* So wird ein echter **autoregressiver Forecast** erzeugt

**🚀 Ideen für die Erweiterung der Forecast-Seite**

Hier ein paar sinnvolle nächste Schritte:

1. **Modellvergleich sichtbar machen**
   * Ein Tab oder Dropdown, das sowohl die Ridge- als auch die Random-Forest-Prognose nebeneinander zeigt
   * Option: Cross-Validation-Fehler als Balkendiagramm anzeigen
2. **Unsicherheiten visualisieren**
   * Konfidenzintervalle oder Prediction-Intervals mit Bootstrap-Sampling
   * Besonders bei Random Forest kann man die Streuung der Einzelbäume nutzen
3. **Mehr Szenarien**
   * Statt nur ±2% beim Energieverbrauch:
     + "Business-as-usual" (keine Veränderung)
     + "Dekarbonisierung" (PEC sinkt -3 %/Jahr, FEC sinkt -1 %)
     + "Worst Case" (PEC steigt +3 %, mehr Verluste)
4. **Hyperparameter-Tuning direkt in der App**
   * Sliders für Ridge-Alpha oder RF-Baumanzahl
   * Direktes Neutrainieren & Vergleich der Ergebnisse
5. **Batch-Vorhersage**
   * Button: „Alle Länder vorhersagen“
   * Ausgabe: Heatmap oder CSV aller Länder/Forecast-Jahre
6. **Explainable AI**
   * Feature-Importances (z. B. von Random Forest)
   * Shapley Values zur Erklärung, warum ein Land hohe Verluste hat

**💡 Nächste Schritte**

Falls du willst, können wir:

* 📄 eine **kurze Dokumentation** (z. B. ML\_README.md) erstellen mit dieser Theorie + Erklärung der Features
* 🚀 eine **Batch-Forecast-Option** hinzufügen, damit du für alle Länder in einem Rutsch eine CSV bekommst
* 🔧 einen kleinen **Hyperparameter-Tuning-Tab** bauen, um z. B. n\_estimators oder alpha auszuprobieren

Möchtest du, dass ich gleich ein kurzes ML\_README.md mit diesen theoretischen Hintergründen erstelle (damit du es in dein Projekt legen kannst)?  
Und soll ich als erste sinnvolle Erweiterung gleich **Batch-Forecast + Download** in ml\_page.py einbauen? (Das wäre eine sehr nützliche Ergänzung für Analysen.)

**You said:**

ja, wir können mit allen deinen Vorschläge weitermachen, aber lieber erst morgen, wenn ich heute Nacht geschlafen habe, Danke, bis morgen!

**ChatGPT said:**

Sehr gerne – das klingt nach einem guten Plan! 😄  
Dann machen wir morgen in Ruhe weiter mit:

✅ **ML\_README.md** – saubere theoretische Dokumentation für dein Projekt  
✅ **Batch-Forecast + CSV-Download** in ml\_page.py  
✅ Optional: erste Schritte zu Hyperparameter-Tuning & Szenarien

Schlaf gut und bis morgen – dann bringen wir dein Projekt noch einen großen Schritt weiter! 🌙💻

Top of Form

Bottom of Form

ChatGPT can make mistakes. Check important info. See Cookie Preferences.

Liebe Mima,

Ich habe seitdem ich meine Unterlagen an dich vor 9 Wochen eingereicht habe, die von meiner Seite vollständig waren, sehe ich immer noch keine eingegangene Zahlung. Ich würde gerne daher die Stellungnahme von Frau Lutscher dazu bekommen, die wahrscheinlich die Aufgaben von Frau Thumm neu übernommen hat, wie es passieren kann, auch wenn Sie vor mehr als 13 Tagen als “die AOs werden nach Durchsicht an die Kasse zur Auszahlung weitergegeben” geschrieben hatte, immer noch kein Zahlungseingang erfolgt ist.

Ich wollte hier an allen beteiligten, die für die Erstellung der Lehraufträge zuständig sind, anfragen, wie man im Anhang sehen kann, dass ich die Lehraufträge für WS25/26 bekommen habe, die Unstimmigkeiten haben. Ich habe die Lehraufträge ohne die vollständigen Namen der Veranstaltungen bekommen, bei den ich den Umfang meiner Lehrverantwortung weder als Stundenzahl noch als Namen nicht nachvollziehen kann (keine ASHE 3a und keine ASHE2a Bezeichnungen dabei). Alle haben nur “2” als Vermerk für Stunden. Ich sollte als Gesamtverantwortliche 2 UE pro Woche für ASHE 3a Übungen haben, aber für die Vorlesungen von ASHE 3a sollte ich insgesamt nur einmalig 4 UE (keine Kompaktseminar, daher Vermerk wäre falsch) und für ASHE 2a sollte ich dieses Semester eigentlich keine Aufträge bekommen. Können die Zuständigen bitte diese Inhalte korrigieren?

Außerdem habe ich immer noch die Stundenhonorare, die ich seit 2-3 Jahren bereits erhalte, als Angaben auf den aktuell erstellten Lehraufträgen (Übungen pro UE 35 E und Seminar 55 E). @Liebe Mima, daher deine Aussage über die nochmalige Überarbeitung von den an die ZV geschickte Lehrauftragsabrechnungen aufgrund der “Erhöhung der Honorare als Stundenlohn” sehe ich als Begründung auf den neu erteilten Lehraufträgen nicht. Es gibt daher nur 2 Möglichkeiten, entweder sind die neuen Lehraufträge, ohne Lohnerhöhung zu berücksichtigen, falsch erstellt, oder deine Begründung für die Verspätung war nicht der Fall, die von Frau Lutschers Seite als ‘aufgrund der Unstimmigkeiten bestand Klärungsbedarf, daher wurden die Dokumente zurückgeschickt” angegeben wurde und hatten eigentlich anderen Gründe, die ich gerne erfahren würde.

Ich würde künftig höflichst alle Beteiligten für die Erstellungen- und Bearbeitungen der Lehraufträge zuständig sind, darum bitten, mich als direkt betroffene Person bei den Problemen mit Lehraufträgen rechtzeitig zu benachrichtigen, sowie mich von cc. zu nicht zu entfernen (was leider Frau Lutscher gemacht hatte als sie am 28.08. nur Frau Batalovic beantwortete), damit ich über die Lage rechtzeitig Bescheid wissen kann.