UCLA Samueli Computer Science

RevMax: Revenue-Maximizing Recommendation System Competition

EAAI-26 Model AI Assignment Track

Fang Sun, Paul Zhang, Pranav Subbaraman, Yizhou Sun
UCLA Computer Science Dept.
fts@cs.ucla.edu

UCLA Samueli Computer Science

Motivation: Beyond Accuracy

Traditional Recommender System assignments often focus on:

- Static datasets (MovieLens, Netflix Prize).
- Accuracy metrics (RMSE, Precision).
- Offline evaluation only.

RevMax shifts the focus to Real-World Dynamics:

- **Business Metric:** Revenue maximization (Price $\times$ Purchase Probability).
- **Dynamic Environment:** Multi-iteration simulation where models *must* adapt.
- **Exploration vs. Exploitation:** Balancing immediate profit with learning user preferences.

UCLA Samueli Computer Science

The RevMax Platform

Built on the **Sim4Rec** simulation framework.

1. **Simulation Loop:**
    - Recommender proposes Top-K items.
    - Simulated Users respond (purchase/ignore).
    - System calculates Revenue.
    - Recommender receives feedback and updates model.
2. **Key Challenge:**
    - Optimizing for **Cumulative Discounted Revenue** over time.
    - Handling "Cold Start" for new users/items introduced during simulation.

UCLA Samueli Computer Science

Checkpoint 1: Content-Based Filtering

"Users like items with similar properties."

- **Focus:**
  - **Feature Engineering:** Extracting value from static user/item attributes (price, category, demographics).
  - **Cold Start Problem:** Handling new items or users with zero interaction history—a critical real-world challenge.
- **Technologies:** Scikit-learn, Random Forests, Gradient Boosting (XGBoost/LightGBM).
- **Learning Goal:** Mastering tabular data processing and traditional supervised learning pipelines.

UCLA Samueli Computer Science

Checkpoint 2: Sequence-Based Modeling

"What did the user interact with *just now* matters most."

- **Focus:**
  - **Temporal Dynamics:** Capturing evolving user interests and short-term intent.
  - **Sequential Dependencies:** Modeling the order of actions (e.g., *view phone* $\rightarrow$ *view case* $\rightarrow$ *purchase*).
- **Technologies:** PyTorch, RNNs/LSTMs, Transformers (SASRec/BERT4Rec).
- **Learning Goal:** Implementing deep learning architectures for sequential data.

UCLA Samueli Computer Science

Checkpoint 3: Graph-Based Approaches

"Users connected to similar items share latent preferences."

- **Focus:**
  - **Bipartite Graphs:** Modeling the explicit User-Item interaction network.
  - **Link Prediction:** Framing recommendation as predicting missing edges in the graph.
  - **High-Order Connectivity:** capturing collaborative signals beyond direct neighbors.
- **Technologies:** Graph Neural Networks (GCN, GraphSAGE), PyTorch Geometric.
- **Learning Goal:** Understanding geometric deep learning and structural data representations.

UCLA Samueli Computer Science

Technical Stack & Student Experience

Designed for **CS145: Introduction to Data Mining** at UCLA.

- **Language:** Python 3.10+
- **Infrastructure:**
  - **PySpark:** Exposure to distributed computing concepts.
  - **PyTorch:** Deep learning implementation.
  - **uv:** Modern, fast Python package management.
- **Workflow:**
  - Students implement `fit()` and `predict()` methods.
  - Provided `Analyzer` tool visualizes learning curves and revenue accumulation.

UCLA Samueli Computer Science

Data Analysis: Understanding the Simulation

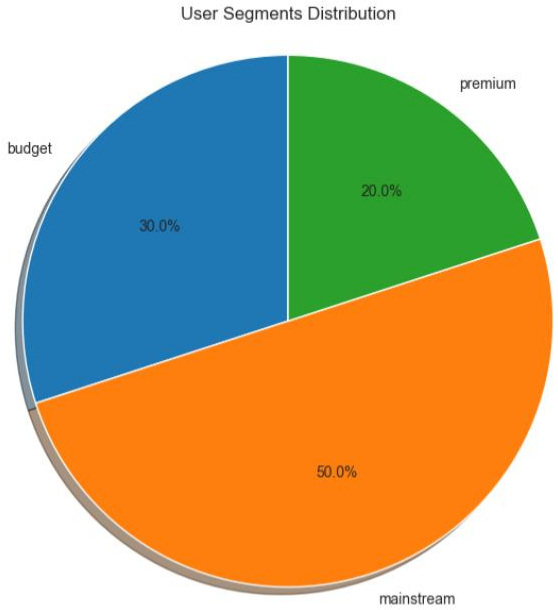
Students begin by exploring the generated synthetic data.

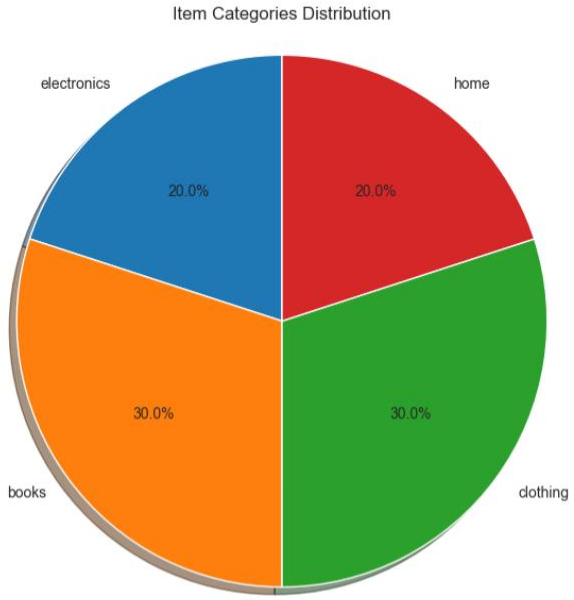
- **User Segments:** Distributions of user types (e.g., price-sensitive vs. quality-focused).
- **Item Categories:** Product hierarchies and price distributions.
- **Interaction Patterns:** Heatmaps of user-item engagement.

UCLA Samueli Computer Science

Generated Data: User & Item Distributions

Left: User Segments | Right: Item Categories





- Visualizing these distributions helps students identify bias and class imbalance.
- Encourages segment-specific strategies.

UCLA Samueli Computer Science

Generated Data: Price & Interactions

Left: Price Distribution | Right: Interaction Heatmap

- **Price:** Essential for optimizing Revenue (not just CTR).
- **Heatmap:** Reveals which user segments prefer which categories.

UCLA Samueli Computer Science

Baseline Models To Start With

The platform includes standard baselines for comparison:

1. **Random:** Stochastic selection (lower bound).
2. **Popularity:** Recommends most purchased items.
3. **SVM/Linear:** Simple supervised learning.

UCLA Samueli Computer Science

Evaluation & The Leaderboard

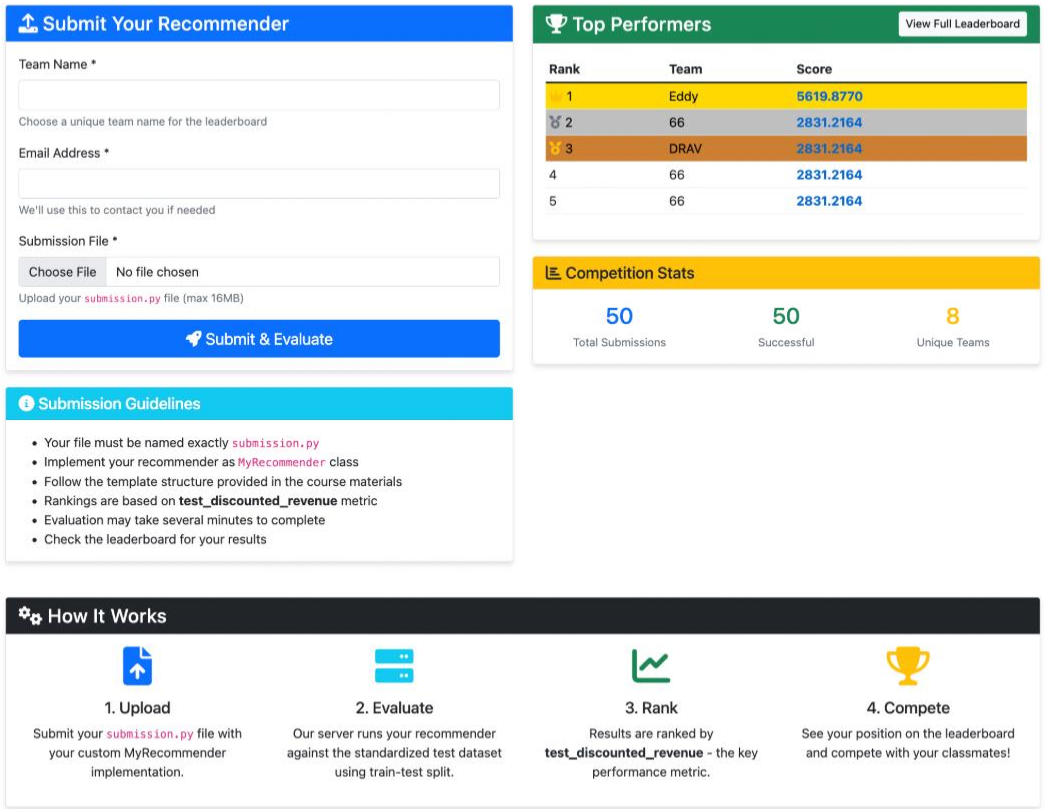
Competition drives engagement.

- **Primary Metric:** Discounted Revenue (Price weighted by rank).
- **Secondary Metrics:** NDCG, Precision@K, Hit Rate.
- **Robustness:**
  - **Hidden Test Environments:** Submissions evaluated on unseen random seeds.
  - Prevents overfitting to the training simulation parameters.

UCLA Samueli Computer Science

Live Leaderboard

- Real-time feedback for students.
- Gamified experience encourages iterative improvement.
- Supports multiple baselines (Random, Popularity) for reference.



UCLA Samueli Computer Science

Classroom Integration

Flexible & Modular:

- **Duration:** Can be run as a semester-long project or in short 2-week modules.
- **Difficulty:** Checkpoints can be assigned individually.
  - *Undergrad:* Focus on Content-based & basic Sequences.
  - *Grad:* Full Deep Learning & Graph implementations.

Provided Resources:

- Detailed Directives for each checkpoint.
- Baseline implementations (Starter code).
- Visualization tools for error analysis.
- Auto-grading scripts.

UCLA Samueli Computer Science

Learning Outcomes

By the end of RevMax, students can:

1. **Translate Business Goals to Code:** Optimizing for *revenue* instead of just *clicks*.
2. **Handle Dynamic Data:** Implementing online learning and frequent model updates.
3. **Implement Modern Architectures:** From Logistic Regression to Transformers and GNNs.
4. **Manage Trade-offs:** Latency vs. Accuracy, Exploration vs. Exploitation.

UCLA Samueli Computer Science

Thank You!

RevMax is available for the EAAI Community.

- **Website:** https://modelai.gettysburg.edu/2026/revmax/
- **Code:** https://github.com/FrancoTSolis/RevMax
- **Contact:** fts@cs.ucla.edu

Ready to maximize revenue in your classroom?