Bayesian Based Bandit Solution For Music Recommendation

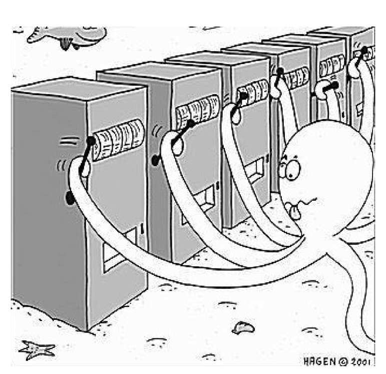
Barış Can Esmer, Beyza Gül Gürbüz, H. Esra Aydemir, Gökçe Uludoğan, Advisor: A. Taylan Cemgil

Bogazici University, Computer Engineering, CMPE547 - Bayesian Statistics

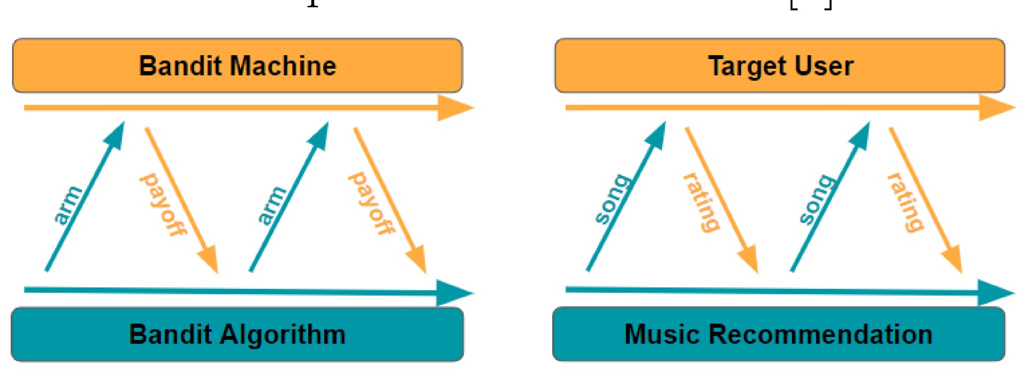
Introduction

Most of the music recommenders recommend songs with the highest user ratings and do not explore the user preferences, so it fails to recommend new rising songs or to new users. However, a good recommender must balance exploring user preferences and exploiting user ratings. In this work, music recommendation is formulated as a multi-armed bandit problem.

Multi-Armed Bandit



A multi-armed bandit problem is a sequential allocation problem defined by a set of actions. At each time step, a unit resource is allocated to an action and some observable payoff is obtained. The goal is to maximize the total payoff obtained in a sequence of allocations. [1]



Dataset

MSD genre dataset is used for song contents[2]. We used a sample containing 6568 songs.
Features of songs are as follows:

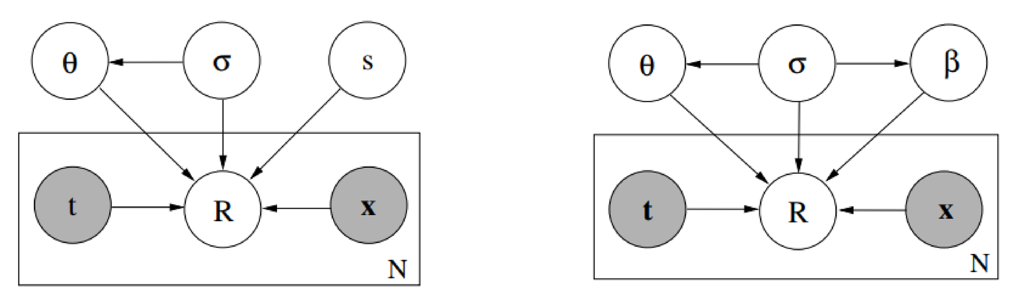
- duration
- key
- mode
- genre
- loudness
- tempo
- time signature
- avg timbres
- var timbres

Model

Personalized, interactive, audio content and novelty based user rating model:

$$U = U_c U_n = \theta^T x (1 - e^{(1-t/s)})$$

where θ represent the user preferences, x the song features, t the time elapsed since the last time the song is listened and finally, s the recovery speed of the novelty.



(a) Exact Model (b) Approximate Model

Approximate Model

$$R \mid x, t, \theta, \beta, \sigma^2 \sim \mathcal{N}(\theta' x \beta' t, \sigma^2)$$
$$\theta \mid \sigma^2 \sim \mathcal{N}(\mu_{\theta 0}, \sigma^2 D_0)$$
$$\beta \mid \sigma^2 \sim \mathcal{N}(\mu_{\beta 0}, \sigma^2 E_0)$$
$$\tau = 1/\sigma^2 \sim \mathcal{G}(a_0, b_0)$$

Methods

Bayes-UCB

The posterior distribution of the user parameters, $\Omega = \{\theta, s\}$, can be given as

$$P(\Omega \mid D_l) \propto P(\Omega) P(D_l \mid \Omega) \quad (1)$$

$$\lambda_k^l = p(\mathcal{U}_k \mid \mathcal{D}_l) = \int p(\mathcal{U}_k \mid \Omega) p(\Omega \mid \mathcal{D}_l) d\Omega \quad (2)$$

where $\mathcal{D}_l = \{(x_i, t_i, r_i)\}_{i=1}^{l}$. Bayes-UCB then recommends the song k^* that satisfies:

$$k^* = arg \max_{k=1...|S|} \mathcal{Q}(\alpha, \lambda_k^l)$$

where $\mathcal{Q}$ is the quantile (generalized inverse) function.
To make the algorithm more responsive, a highly efficient variational inference algorithm is proposed for approximating the inference procedure.

Methods

Variational Bayes-UCB

Following the convention of mean-field approximation, we assume that the joint posterior distribution factorizes as follows:

$$p(\Omega|\mathcal{D}) = p(\theta, \beta, \tau|\mathcal{D}) \approx q(\theta, \beta, \tau) = q(\theta) q(\beta) q(\tau)$$

Because of the choice of conjugate priors each factor distribution q(θ), q(β) and q(τ) take the same parametric forms as the prior distributions. Specifically,

$$q(\theta) \propto exp(-\frac{1}{2}\theta^T \Lambda_{\theta N} \theta + \eta_{\theta N}^T \theta)$$
$$q(\beta) \propto exp(-\frac{1}{2}\beta^T \Lambda_{\beta N} \beta + \eta_{\beta N}^T \beta)$$
$$q(\tau) \propto \tau^{a_N - 1} exp(-b_N \tau)$$

For optimization, we use the coordinate descent method to minimize

$$KL(\, p(\theta, \beta, \tau|\mathcal{D}) \,||\, q(\theta) q(\beta) q(\tau) \,)$$

Finally, since $q(\theta)$ and $q(\beta)$ are normal distributions and linear combination of normal random variables is again a normal random variable, we obtain:

$$p(\theta^T x | x, t, \mathcal{D}) \approx \mathcal{N}(x^T \Lambda_{\theta N}^{-1} \eta_{\theta N}, x^T \Lambda_{\theta N}^{-1} x)$$
$$p(\beta^T t | x, t, \mathcal{D}) \approx \mathcal{N}(t^T \Lambda_{\beta N}^{-1} \eta_{\beta N}, x^T \Lambda_{\beta N}^{-1} x)$$

and posterior distribution in Equation 2 can be calculated as

$$p(\mathcal{U} \mid x, t, \mathcal{D}) = p(\theta^T x \beta^T t \mid x, t, \mathcal{D}) =$$
$$\int p(\theta^T x = a \mid x, t, \mathcal{D}) p(\beta^T t = \frac{\mathcal{U}}{a} \mid x, t, \mathcal{D}) da$$

Greedy & ε-greedy

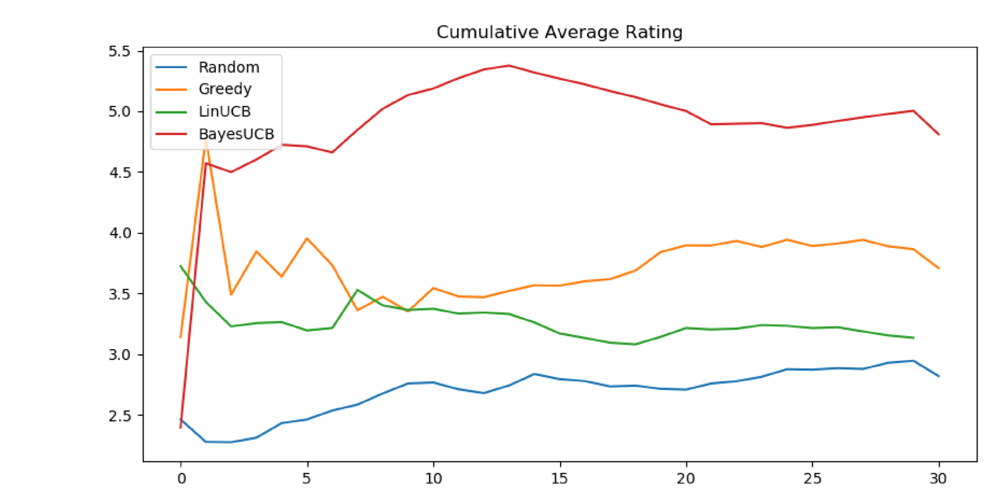
- Recommends song with highest expected rating.
- Pure exploitation (Exploration with ε in ε-greedy)
- L-BFGS-B optimization with minimum MSE.

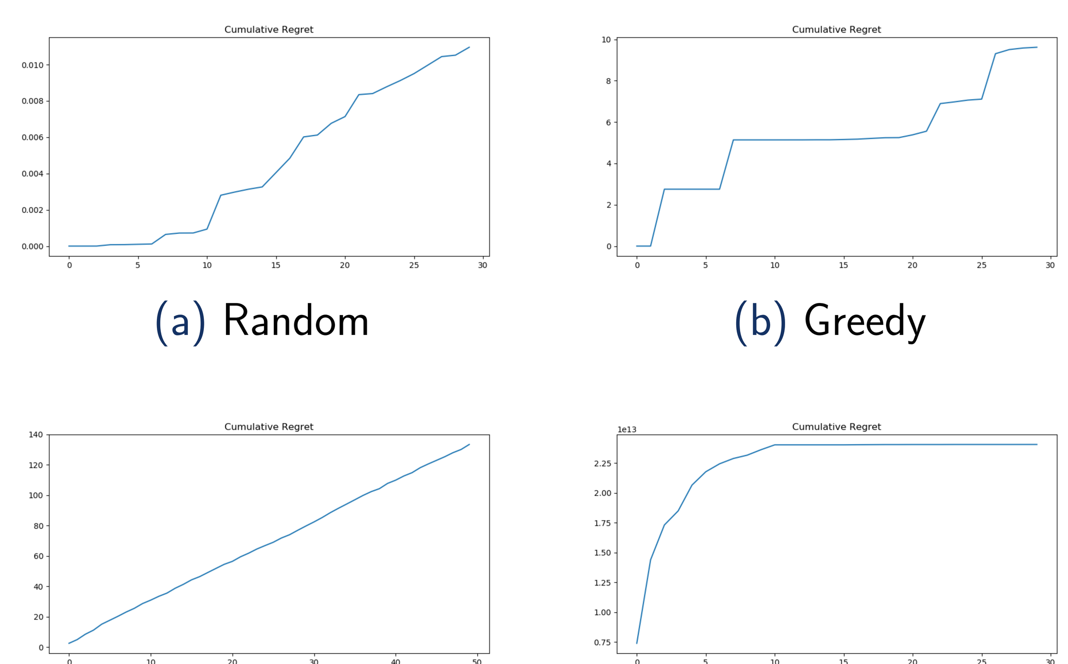
Lin-UCB

- Assumes expected rating is a linear function of the features.
- Ridge regression with upper confidence bound.
- Balances exploration and exploitation.

Experiments

We have simulated user actions with respect to the information from our dataset and used this simulation for evaluation purposes.





(a) Random (b) Greedy (c) Lin-UCB (d) Bayes-UCB

Conclusion

We have specifically investigated Bayes-UCB method for multi-armed bandit on music recommendation problem and compared it with other methods. We observed that Bayes-UCB surpasses other methods in terms of cumulative rating and regret.

References

[1] Sébastien Bubeck, Nicolo Cesa-Bianchi, et al.
Regret analysis of stochastic and nonstochastic multi-armed bandit problems.

[2] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere.
The million song dataset.

[3] Lihong Li, Wei Chu, John Langford, and Robert E Schapire.
A contextual-bandit approach to personalized news article recommendation.

[4] Xinxi Wang, Yi Wang, David Hsu, and Ye Wang.
Exploration in interactive personalized music recommendation: a reinforcement learning approach.