CS 181 Machine Learning
Practical 4 Report, Team *la Dernière Dame M*

(Jeremiah) Zhe Liu[1], (Vivian) Wenwan Yang[2], and Jing Wen[1]

[1]Department of Biostatistics, Harvard School of Public Health
[2]Department of Computational Science and Engineering, SEAS

May 6, 2015

1 Problem Description

Set in a *Flappy Bird*-type game *Swingy Monkey*, our current learning goal is to estimate an optimal policy $\pi : \mathcal{S} \to \mathcal{A}$ such that the expectation of reward function $f : \mathcal{S} \times \mathcal{A} \to \mathcal{R}$ is maximized, i.e. we aim to identify a π^* such that

$$\pi^* = \arg\max_{\pi} E\Big(f(s, \pi(s))|s\Big)$$

State Space Action Space
Empirical Goal: Score Reward

2 Method

2.1 Rationale on Model Choice

2.1.1 State Reduction and Discretization

2.1.2 Exploration/Exploitation Parameters

Learning rate
ϵ-greedy

3 Result

3.1 State Exploration

3.2 Convergence Behavior

4 Discussion & Possible Directions

Reference

1. Ricci F, Rokach L, Shapira B et al. (2010) **Recommender Systems Handbook**. *Springer*.

2. Koren Y, Bell R, Volinsky C. (2009) **Matrix factorization techniques for recommender systems**. *IEEE Computer* Aug 2009, 42-49.

3. Srebro N, Jaakkola T.(2003) **Weighted low-rank approximations**. *Proceedings of the Twentieth International Conference* 720727.

4. R Salakhutdinov, A Mnih. (2008) **Probabilistic Matrix Factorization**. *Advances in Neural Information Processing Systems* Vol. 20

5. Koren, Y. (2008) **Factorization Meets the Neighborhood: a Multifaceted Collaborative Filtering Model**, *Proc. 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.