ID de Contribution: **214** Type: **Invited talk**

Global convergence of stochastic gradient bandits for any learning rates

We provide a new understanding of the stochastic gradient bandit algorithm by showing that it converges to a globally optimal policy almost surely using **any** constant learning rate. This result demonstrates that the stochastic gradient algorithm continues to balance exploration and exploitation appropriately even in scenarios where standard smoothness and noise control assumptions break down. The proofs are based on novel findings about action sampling rates and the relationship between cumulative progress and noise, and extend the current understanding of how simple stochastic gradient methods behave in bandit settings.

Author: MEI, Jincheng (Google DeepMind)

Orateur: MEI, Jincheng (Google DeepMind)

Classification de Session: ML

Classification de thématique: Machine learning