الجمهورية الشعبية الديمقراطية الجزائرية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي 08 ماي 5491 بسيدي بلعباس
École Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès

THESIS

To obtain the diploma of **Engineer**
Field: **Computer Science**
Specialty: **Information Systems and Web (SIW)**

Theme

Comparing Multi-Armed Bandit Algorithms (MABs): In-Memory vs. In-Database implementation for a recommendation system

---

Presented by:
YOUSFI Ikram Oumelkheir

Submission Date: **September 24, 2024**
In front of the jury composed of:

Ms. BENCHERIF Khayra	President
Ms. KLOUCHE Badia	Examiner
Ms. ELOUALI Nadia	Supervisor
Ms. AMER-YAHIA Sihem	Co-Supervisor
Mr. BOUAROUR Nassim	Co-Supervisor

Academic Year : 2023/2024

Abstract

Recommendation systems have become essential components of modern applications, enhancing user experiences with personalized suggestions through custom filtering and prioritization. However, traditional systems often struggle with scalability, particularly when faced with increasing data volumes and the demand for real-time personalization.

Our study focuses on developing and training MAB models using SQL queries within the database, a method known as in-Database Machine Learning (in-DB ML). This approach offers significant advantages, such as reduced latency and minimized data movement, particularly in scenarios involving small datasets or models. Notably, in small-scale studies, SQLite demonstrated faster runtime performance compared to Python, highlighting the potential of in-DB ML. This shift towards in-DB ML signifies a broader trend in the evolving role of databases—not just as repositories for data storage and retrieval, but as integral platforms for executing sophisticated analytical tasks directly at the data source.

Keywords— In-Database Machine Learning, Multi Armed Bandits, Recommendation System, SQLite, MABs

Résumé

Les systèmes de recommandation sont devenus des partenaires essentiels dans les applications modernes, aidant les utilisateurs en personnalisant les options grâce à un filtrage et une mise en avant sur mesure. Cependant, les systèmes de recommandation traditionnels rencontrent souvent des problèmes de scalabilité, ayant du mal à gérer l'augmentation des volumes de données et la nécessité de personnalisation en temps réel.

Dans notre recherche, nous utiliserons les bandits manchots (Multi-Armed Bandits), qui sont des modèles robustes spécialement conçus pour la prise de décision dans des environnements incertains, dans le but d'améliorer davantage le processus de recommandation.

Notre recherche se concentre plutôt sur le développement et l'entraînement des modèles MAB au sein d'une base de données avec des requêtes SQL. Bien que cette méthodologie de machine learning en base de données (in-DB ML) présente plusieurs avantages en termes de réduction de la latence et des mouvements de données, elle peut être bénéfique pour des petits ensembles de données ou des petits modèles. En effet, le temps d'exécution de SQLite est beaucoup plus rapide que celui de Python dans les études à petite échelle, révélant ainsi le potentiel de l'in-DB ML. Cette tendance vers le ML en base de données souligne comment le rôle des bases de données est en train de changer : elles ne se contentent plus de stocker et de récupérer des données, mais deviennent centrales pour débloquer des tâches analytiques sophistiquées situées directement à la source des données.

Mots clés— In-Database Machine Learning, Multi Armed Bandits, Recommendation system, SQLite, MABs

الملخص

لقد تطورت أنظمة التوصية لتصبح شركاء أساسيين في التطبيقات الحديثة، حيث تساعد المستخدمين من خلال تخصيص الخيارات عبر التصفية المخصصة وإبراز المحتوى المناسب. ومع ذلك، غالباً ما تواجه أنظمة التوصية التقليدية تحديات في التوسع، مما يجعل من الصعب التعامل مع كميات البيانات المتزايدة والحاجة إلى التخصيص في الوقت الحقيقي.

في بحثنا، سنستخدم نماذج الـ Multi-Armed Bandits (MABs) ، وهي نماذج قوية مصممة خصيصًا لاتخاذ القرارات في البيئات غير المؤكدة، بهدف تعزيز عملية التوصية. تشمل الفوائد العملية لتعلم الآلة داخل قواعد البيانات انخفاض زمن الوصول في توليد الرؤى، وتحسين حوكمة البيانات، وتبسيط البنية التحتية لتدفقات العمل التحليلية. والأهم من ذلك، أن هذا التكامل يجعل عمليات تعلم الآلة القابلة للتوسع على مجموعات بيانات كبيرة ممكنة في الأنظمة المؤسسية، مما يجعلها ذات صلة بحالات الاستخدام في القطاعات مثل المالية والرعاية الصحية والتجزئة، حيث تكون خصوصية البيانات وسرعة معالجة البيانات ضرورية.

يركز بحثنا بدلاً من ذلك على تطوير وتدريب نماذج الـ MAB داخل قاعدة بيانات باستخدام استعلامات SQL . وبينما يجلب هذا النهج في تعلم الآلة داخل قواعد البيانات In-DB ML العديد من الفوائد مثل تقليل زمن الانتظار وحركة البيانات، يمكن أن يكون ذلك مفيدًا في مجموعات البيانات الصغيرة أو النماذج الصغيرة. في الواقع، أظهرت الدراسات الصغيرة أن وقت تشغيل SQLite أسرع بكثير من Python ، مما يكشف عن إمكانيات تعلم الآلة داخل قواعد البيانات. ويؤكد هذا الاتجاه نحو In-DB ML على كيف يتغير دور قواعد البيانات: لم تعد مجرد تخزين واسترجاع البيانات، بل أصبحت محورًا لإجراء المهام التحليلية المتقدمة مباشرة في مصدر البيانات.

الكلمات المفتاحية: تعلم الآلة داخل قاعدة البيانات, النطاقات متعددة الأذرع , نظام ,التوصية SQLite , MABs .