



جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

École Nationale Supérieure d'Informatique et d'Analyse des Systèmes
Centre d'Études Doctorales en Sciences des Technologies de l'Information et de l'Ingénieur

AVIS DE SOUTENANCE DE THÈSE DE DOCTORAT

Madame Fatima Zahra LAHLOU

soutiendra publiquement sa thèse de Doctorat en Informatique

le Lundi 30 Décembre 2019 à 10h00 à l'Amphi I à l'ENSIAS

Intitulé de la thèse :

REVIEW BASED CONTEXT AWARE RECOMMENDER SYSTEMS

Devant le Jury composé de :

Président :

Pr. Bouchaïb BOUNABAT, PES, ENSIAS, Université Mohammed V de Rabat

Directeur de thèse :

Pr. Ismail KASSOU, PES, ENSIAS, Université Mohammed V de Rabat

Co-encadrante de thèse :

Pr. Houda BENBRAHIM, PH, ENSIAS, Université Mohammed V de Rabat

Rapporteurs :

Pr. Saïd SLAOUI, PES, FSR, Université Mohammed V de Rabat

Pr. Omar EL BEQQALI, PES, FSDM, Université Sidi Mohamed Ben Abdellah, Fès

Pr. Bouchra EL ASRI, PES, ENSIAS, Université Mohammed V de Rabat

Examinatrice :

Pr. Sanaa EL FKIH, PH, ENSIAS, Université Mohammed V de Rabat

REVIEW BASED CONTEXT AWARE RECOMMENDER SYSTEMS

Abstract: Recommender Systems (RS) are systems that assist users during their navigation by suggesting relevant items to them. While traditional RS focus only on users and items for predicting which items are relevant, Context Aware Recommender Systems (CARS) consider, in addition to users and items, contextual information. Getting contextual information is a first challenge for building any context aware recommender system. While users' reviews often contain contextual information, we address throughout this dissertation the research issue of how to efficiently leverage context available in users' reviews to produce context aware recommendations.

During this dissertation, in a first time, we explore the use of text classification techniques for inferring context from reviews, and try to identify the best text preprocessing and representation settings suitable for this task. To achieve this, we conduct experiments and compare different text representation settings and classification algorithms in order to determine those that are appropriate for the context inference problem. Result demonstrate promising classification accuracy, but not enough to build context aware recommender on it.

After that, we adopt a different approach, and present a new system architecture that take the whole review as input. This architecture, that we call Review Aware Recommender Systems, aims to bring out context from reviews and incorporate it in the recommender algorithm simultaneously, automatically, and without requiring specific feature engineering. An empirical evaluation shows that using extracted contexts throughout this architecture significantly improve recommendations. We consider this as new and important finding in CARS research.

Furthermore, we develop a new CARS algorithm, that is tailored to such situations where the context is derived from text, and that we call Textual Context Aware Factorization Machines. We implement our new algorithm in the TensorFlow framework (Abadi et al., 2016), and demonstrate throughout conducted

experiments that our proposed algorithm, with appropriate settings, outperforms the state of art CARS algorithm, Context Aware Factorization Machines (Rendle et al., 2011).

Keywords: Recommender Systems; Context Aware Recommender Systems; Factorization Machines; Text Mining; Text Representation; Text Classification; Machine Learning.

ملخص

أنظمة التوصية (أو أنظمة الاقتراح) هي فرع من أنظمة تصفية المعلومات التي تهدف إلى مساعدة المستخدمين أثناء التصفح عن طريق اقتراح منتجات يقع عليها إعجاب المستخدم. بينما تبني أنظمة التوصية التوقعات انطلاقاً من معلومات حول المستخدمين والمنتجات، تستعمل أنظمة التوصية الحساسة للسياق معلومات إضافية حول سياق المستخدمين.

تعد مرحلة الحصول على معلومات عن السياق أول خطوة لإنشاء نظام توصية حساس للسياق. ولأن هذه المعلومات كثيراً ما تتواجد ضمن تعليقات المستخدمين، ندرس خلال هذه الأطروحة إشكالية الاستخدام الفعال للمعلومات السياقية الواردة في تعليقات المستخدم لإنتاج توصية حساسة للسياق.

خلال هذه الأطروحة، نستكشف أولاً مدى دقة إمكانية استنتاج السياق الوارد ضمن تعليقات المستخدم باستعمال تقنيات تصنيف النص، وما هي الإعدادات الأكثر ملائمة لتمثيل ومعالجة النصوص من أجل هذا الغرض. لتحقيق ذلك، نقوم بإجراء تجارب مختلفة نقارن من خلالها إعدادات تمثيل النص وخوارزميات التصنيف من أجل تحديد تلك المناسبة لمشكلة استنتاج السياق. تظهر النتائج دقة تصنيف واعدة، ولكنها ليست كافية للاعتماد عليها في نظام توصية حساسة للسياق.

بعد ذلك، نعتمد نهجاً مختلفاً، ونقدم بنية تهدف إلى استخراج السياق تلقائياً من تعليقات المستخدم وإدراجه في خوارزمية التوصية في وقت واحد. يظهر تقييم تجريبي أن استخدام السياقات المستخرجة باستعمال هذه البنية يحسن التوصيات بشكل واضح. يمكن اعتبار هذه النتائج اكتشافاً جديداً ومهماً في أبحاث أنظمة التوصية الحساسة للسياق.

علاوة على ذلك، نقوم بتطوير خوارزمية جديدة لنظام توصية حساس للسياق، مصممة خصيصاً لمثل هذه الحالة حيث يتم اشتقاق السياق من النص. أظهرت التجارب التي أجريت أن خوارزمتنا المقترحة، مع إعدادات المناسبة، تتفوق على الخوارزمية المرجعية لأنظمة التوصية الحساسة للسياق.

الكلمات الرئيسية

أنظمة التوصية؛ أنظمة التوصية الحساسة للسياق؛ آلات التعميل؛ التنقيب في النصوص؛ تمثيل النصوص؛ تصنيف النصوص؛ التعلم الآلي.