

الجزائرية الديمقراطية الشعبية الجمهورية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



École Supérieure en Informatique Sidi Bel Abbés - 8 mai 1945

Mémoire de fin d'étude

En vue de l'obtention du diplôme : Master

Filière : Informatique

Spécialité : Système d'Information et Web (SIW)

Thème

Time series representation

Présenté par : Ghiles Meddour

Soutenu le : xx/10/2020

Devant le jury composé de :

Mr. xxx xxx	Docteur	Président
Mr. xxx xxx	Docteur	Encadrant
Mr. xxx xxx	Professeur	Encadrant
Mme. xxx xxx	Docteur	Examineur

Année universitaire : 2019/2020

Abstract

Time series have become essential in the data industry and are present in all fields (sensors, IoT, ...). Their exploitation offers a huge potential but presents a great challenge. Indeed, their nature and their large dimensionality make the usual methods of data mining and pattern recognition not very efficient. As a result, several time series representations capable of achieving significant dimensionality reduction without losing important features have been developed. Each method of representation has strengths and weaknesses. In this work, we make a bibliographical review of the different methods of representation of time series and we make a comparison between them. In our review, we are particularly interested in two methods: SAX, a method that is already well established and has many successful applications, and SSTS, a recently proposed but very promising method.

Résumé

Les séries temporelles sont devenues essentielles dans l'industrie des données et sont présentes dans tous les domaines (capteurs, IdO, ...). Leur exploitation offre un potentiel énorme mais représente un grand défi. En effet, leur nature et leur grande dimensionnalité rendent les méthodes habituelles d'exploration de données et de reconnaissance de formes peu efficaces. En conséquence, plusieurs représentations de séries temporelles capables d'atteindre une réduction significative de la dimensionnalité sans perdre des caractéristiques importantes ont été développées. Chaque méthode de représentation présente des forces et des faiblesses. Dans ce travail, nous faisons une revue bibliographique des différentes méthodes de représentation des séries temporelles et nous les comparons entre elles. Dans notre revue, nous nous intéressons particulièrement à deux méthodes : SAX, une méthode déjà bien établie et qui a de nombreuses applications réussies, et SSTS, une méthode récemment proposée mais très prometteuse.

الملخص

أصبحت السلاسل الزمنية ضرورية في مجال البيانات وهي موجودة في جميع القطاعات (أجهزة الاستشعار ، إنترنت الأشياء ، ...) . يوفر استغلالهم إمكانات هائلة ولكنه يمثل تحدياً كبيراً. في الواقع ، فإن طبيعتها وأبعادها الكبيرة تجعل الطرق المعتادة لاستخراج البيانات والتعرف على الأنماط غير فعالة للغاية. نتيجة لذلك ، تم تطوير العديد من تمثيلات السلاسل الزمنية القادرة على تحقيق تقليل كبير في الأبعاد دون فقدان الميزات المهمة. كل طريقة من طرق التمثيل لها نقاط قوة وضعف. في هذا العمل ، نجري مراجعة بليوغرافية لمختلف طرق تمثيل السلاسل الزمنية ونجري مقارنة بينها. في مراجعتنا ، نحن مهتمون بشكل خاص بطريقتين: SAX، وهي طريقة راسخة بالفعل ولها العديد من التطبيقات الناجحة ، و SSTS ، وهي طريقة مقترحة مؤخراً ولكنها واعدة للغاية.