وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم البرمجيات

تمييز جودة الحليب باستخدام التعلم الآلي

رسالة مقدمة

الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل

كجزء من متطلبات نيل شهادة الدبلوم العالي في

البرمجيات

من قبل

نور عبد الخالق فاضل عبد

إشراف

أ.م. بيداء سليمان بهنام

1445هـ 2024م

الملخص

تعد جودة الألبان من الأمور ذات الأهمية الكبرى في صناعة الأغذية، إذ تعد الألبان من المصادر الرئيسية للبروتينات، والكالسيوم، والفيتامينات، والمعادن التي يحتاجها جسم الإنسان, الحليب أساس كل منتجات الالبان, فهو يمثل العنصر الأساسي والمهم في حياة الناس، ولاسيما الأطفال، لأنه مصدر لايستغنى عنه لنموهم، وبناء عظامهم، وتقوية أجسامهم.

ويعد تحديد جودة منتج الحليب أمر بالغ الأهمية لغرض المراقبة لتقليل الخسائر والضرر المحتمل، ومن أجل الحصول على كشف دقيق لجودة منتج الحليب, أما الحليب ذو جودة عالية أو ذو جودة منخفضة، وقد تم في هذه الدراسة استخدام خوارزميات التعلم الآلي، وهي خوارزمية آلة متجه الدعم Support Vector Machine (SVM) ، وخوارزمية الجيران الاقرب K-nearest neighbor (KNN)، خوارزمية شجرة القرار Decision Tree (DT) وخوارزمية التعبئة Bagging باستخدام مجموعة بيانات الحليب المأخوذة من مستودع Kaggle. إذ تتكون هذه البيانات من 1059 عينة.

لقد تم تدريب النماذج المقترحة واختبارها بهدف إيجاد الانموذج الأفضل والأكثر دقة للكشف عن جودة الحليب وتم تقييمها باستخدام مقاييس التقييم: الدقة، والضبط، والاستدعاء، f1_score ومصفوفة الارتباك، ووفقاً لنتائج التقييم، تفوقت ثلاثة نماذج: SVM، وKNN، وDT على خوارزمية التعبئة، حيث حصلت على أعلى مستوى لجميع المقاييس بنسبة 100%. وكانت خوارزمية SVM هي الأكثر كفاءة، لأن وقت تنفيذها كان 0.146 ثانية، وهو أقل من النماذج الأخرى.

Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and Mathematics
Department of Software

Distinguishing milk quality using machine learning

A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of High Diploma
in Software

By
Nour Abd AL Khaliq Fadel

Supervised by
Assist. Prof. Baydaa Sulaiman Bahnam

1445 A.H. **2024 A.D.**

Abstract

The quality of dairy is of great importance in the food industry, as dairy is considered one of the main sources of proteins, calcium, vitamins, and minerals that the human body needs. Dairy is found in several products, the most important of which is milk. It represents the basic and important element in people's lives, especially children, because it is an indispensable source for their growth, building their bones, and strengthening their bodies.

Milk is a perishable product. Every gram of milk of poor quality or structure can cause tons of milk to spoil, causing major financial losses. Therefore, the quality and safety standards of this product are essential to ensure the provision of healthy and safe products to consumers.

Determining the quality of the milk product is crucial for the purpose of monitoring to reduce potential losses and damage. In order to obtain an accurate detection of the quality of the milk product, whether high-quality or low-quality milk, machine learning algorithms were used in this thesis, namely the Support Vector Machine (SVM) algorithm, and the K-nearest neighbor (KNN) algorithm, Decision Tree (DT) algorithm and Bagging algorithm using milk dataset taken from Kaggle repository. This data consists of 1059 samples.

The proposed models were trained and tested with the aim of finding the best and most accurate model for detecting milk quality and were evaluated using evaluation metrics: Accuracy, Precision, Recall, F1_score and Confusion matrix. According to the evaluation results, three models: SVM, KNN, and DT outperformed the bagging algorithm, obtaining the highest level of 100% for all metrics. The SVM algorithm was the most efficient because its execution time was 0.146 seconds, which was less than other models.