



**COLLEGE OF GRADUATE STUDIES AND ACADEMIC
ADVANCEMENT**

Faculty of Computer Science and Information Technology

Clustering Technique For Sudanese Female Dress Standard

by

Tasneem Awad Mohammed

A Thesis

Submitted to the College of Graduate Studies and Academic Advancements in
Partial Fulfillment of the Requirement for the Degree of Master of Science in
Computer Science

Supervisor

Dr. Mohamed Hafize Musa

October 2015

Abstract

International standard for women sizing based on sizes taken from different nationalities (Asian, European, American ...) This research tries to answer the question know how American standard suit to Sudanese women dress . To answer this question; data set was collected manually in khartoum area , it were 600 female dress sizes included 6 attribute (shoulder , sleeve , bust ,waist, hips , dress length) , from women of different ages and different Sudanese ethnic groups by using k-mean algorithm to get the best cluster ,and compared with American standards and find out the most related attribute in Sudanese female dresses size. the result of this studies was that ; Sudanese female dress size is not suitable to united states female dress size . The two attributes (bust &waist) of the size XXL in american standard dos not suit to Sudanese female ; in all experiments. Most of Sudanese female sizes found in the range between XS and L suitable to dress . attributes ,bust and shoulder are strongly related attributes in Sudanese female dresses size.

ملخص

المقاسات العالمية للابسة تم بنائها على مقاسات اخذت من اشخاص اسويين واوربيين وامريكان .

يهدف هذا البحث لمعرفة مناسبة هذه المقاسات للأشخاص السودانيين، للوصول لهذا الهدف تم جمع عدد 600 مقاس للفساتين يدويا في منطقة الخرطوم شملت ستة صفات (الكتف، الأكمام والصدر ، والخصر والاردا ف ، وطول الفستان) من النساء من مختلف الأعمار والفئات العرقية السودانية المختلفة باستخدام خوارزمية (k-mean) للحصول على مقاس للفستان يناسب السودانيات وذلك عن طريق مقارنته بمقاس الفستان الامريكي . تمت مقارنه الصفات الصدر والخصر والاردا ف وكانت نتيجة المقارنه هي مقاس الفستان السوداني غير مطابق لمقاس الفستان الامريكي.

في التجربه الثانيه تم اختبار الصفات (الكتف، الأكمام والصدر ، والخصر والاردا ف ، وطول الفستان) لمقاس الفستان السوداني للحصول على اكثر الصفات ارتباطا بمقاس الفستان، وذلك بحذف صفة من الصفات واعادة التجربه ،هناك عشرة مقاسات مختلفه (تشابه او تكرار المقاس يعمل على نقصان عدد المقاسات الى اقل من عشرة) اذا ظهرت عدة مقاسات متشابهه او متكررة هذا دليل على ان الصفة المحذوفه اكثر تأثيرا على مقاس الفستان وكانت نتيجة هذه التجربه كالاتي عند حذف الكتف تكررت المقاسات 4,6,7,9 . وعند حذف الاكمام تكررت المقاسات 6,9 . وعند حذف الصدر تكررت المقاسات 1,2,6,8 . وعند حذف الخصر لا يوجد تكرار للمقاس . وعند حذف الاردا ف لا يوجد تكرار للمقاس . وعند حذف طول الفستان تكرر مقاس 4 . ومن السابق يتضح ان صفة الصدر وصفه الكتف من اكثر الصفات ارتباطا على مقاس الفستان السوداني

1.1 Introduction

Data mining is a knowledge discovery process of extracting previously unknown, actionable information from very large databases. clustering is a data mining (machine learning) technique used to place data elements in to related groups without advance knowledge of the groups definitions[1]. k-means clustering is a data mining/machine learning algorithm used to cluster observations into groups of related observations without any prior knowledge of those relationships. The k-means algorithm is one of the simplest clustering techniques [2].

1.2 Research Problems

The problems of the research is answer the question how American standard suit to Sudanese women dress and witch attribute was the strongest in relating in Sudanese female dresses size . ..

1.3 Research Objectives

Collecting data for comparison with international dress sizes by using k-means clustering algorithm.

1.4 Research Methodology

The data set used in this study'' women dress size'' collected manually in the year 2015 from women dress tailors it were 600 female dress sizes , in khartoum area from women of different ages and different Sudanese ethnic groups .

1.4.1 The Overall Process Of Finding And Interpreting Patterns From Data Involves The Application Of The Following Steps:

Developing an understanding of

- i. the application domain
- ii. the relevant prior knowledge
- iii. the goals of the end-user

Creating a target data set by selecting a data set, or focusing on a subset of variables, or data samples, on which discovery is to be performed.

Data reduction And Projection.

Finding useful features to represent the data depending on the goal of the task

Using dimensionality reduction or transformation methods to reduce the effective number of variables under consideration or to find invariant representations for the data. Choosing The Data Mining Task . Deciding whether the goal of the KDD process is classification, regression, clustering, etc, Choosing The Data Mining Algorithm (s), Selecting method(s) to be used for searching for patterns in the data, Deciding which models and parameters may be appropriate, Deciding which models and parameters may be appropriate, Matching a particular data mining method with the overall criteria of the KDD process, Data Mining . Searching for patterns of interest in a particular representational form or a set of such representations as classification rules or trees, regression, clustering, and so forth, Interpreting mined patterns and Consolidating discovered knowledge [5].

