

Sudan University of Science and Technology
College of Graduate Studies

Speech Recognition Using
Artificial Neural Networks

التعرف على الكلام باستخدام الشبكات العصبية الاصطناعية

By

Wifag Abdallah Elhassan Sid Ahmed

Research Submitted In Fulfillment of the Requirements
For MSc Degree in Computer Science

Supervisor:

Dr. Eltahir Mohammed Hussein

May 2008

Dedication

To my Mother, my Father, Uncle, Brothers, Sisters,

Friends And my Relatives,

To all Whom I love

i

Acknowledgement

My gratitude and thanks to my God. I am grateful to my supervisor Dr. Eltahir Mohammed Hussein, head department of Bio medical engineering in Sudan University, for his invaluable, support and guidance during implementation of this research. My gratitude extended to Dr. Adil Mohammed Ahmed Elsinari who offered his academic experience for his hard help in preparation of this study.

I am very thankful to Sudan University of Science and Technology special College of Computer Science and all staff members of Omdurman Islamic University. Special thanks are due to Engineer Abd Elfatah Mohammed Ahmed. My thanks are due to several of my colleagues and friends for their spiritual support.

Finally, special thanks to my family for their patience and cooperation during the preparation period of this work. I also would like to thank my aunt Dar Elsalam

I gladly acknowledge all those who supported and encouraged me.

The objective of this study is to evaluate the potentiality of using Artificial Neural Networks (ANNs) for Speech Recognition.

The Linear Predictive Code (LPC) was used for the feature extractions of the word used. The speech data (spoken words) has been converted in to voice signals in digital format. MS-Excel package has been used to generate 600 learning pattern. 540 were used to train General Regression Neural Network (GRNN) and Back Propagation Network (BPN) architecture. The reminder 60 patterns were used to test the performance of the trained shell.

The General Regression Neural Network (GRNN) was found to be able to recognize speech patterns and process test patterns with an average error of ± 0.016667 while the standard deviation (STVD) was ± 0.129099 . Otherwise, the average of the BPN was $\pm 1.168 \times 10^{-4}$.

الهدف من هذه الدراسة هو تقييم إمكانية استخدام الشبكات العصبية الاصطناعية لتمييز الكلام.

تم استخلاص السمات للكلمات المستخدمة في هذه الدراسة عن طريق قنوات تشفير التوقع الخطي الذي يعتبر طريقة لتشفير المعلومات في إشارة الكلام ، حيث حولت الكلمة المنطوقة لإشارة صوتية في قالب رقمي.

إستخدم برنامج مايكروسوفت إكسيل لتوليد 600 عينة كنموذج للدراسة. 540 عينة استخدمت لتدريب نموذج الارتداد العام ونموذج الإنتشار الخلفي للشبكات العصبية ، بينما استخدمت 60 عينة لإختبار مقدرة الشبكة المدربة.

أسفرت نتائج الدراسة بأن الشبكات العصبية الاصطناعية ذات الارتداد العام لها القدرة على اتمييز الكلام ومعالجة عينات التدريب بمتوسط خطأ بلغ $0.016667 \pm$ وانحراف معياري بلغ $0.129099 \pm$ بينما بلغ متوسط الخطأ في تدريب الشبكات العصبية ذات الانتشار الخلفي $1.168 \times 10^{-4} \pm$.