كلية التجارة
قسم الإحصاء، الرياضة والتأمين

استخدام الأساليب الإحصائية لاقتراح محافظ الاستثمار فى الأوراق المالية

رسالة مقدمة
للحصول على درجة الماجستير فى الإحصاء التطبيقى

اعداد

منى محمود محمد عبادة

معيدة بكلية التجارة - جامعة عين شمس

تحت إشراف

أ.د.مصطفى جلال مصطفى

قسم الإحصاء، الرياضة والتأمين
كلية التجارة ـ جامعة عين شمس

أ.د. مدحت محمد أحمد عبد العال

رئيس قسم الإحصاء، الرياضة والتأمين
كلية التجارة ـ جامعة عين شمس

2014

بِسْمِ اللهِ الرَّحْمنِ الرَّحِيمِ

۞قَالُوا سُبْحَانَكَ لَا عِلْمَ لَنَا إِلَّا مَا عَلَّمْتَنَا إِنَّكَ أَنْتَ الْعَلِيمُ الْحَكِيمُ۞

صدق الله العظيم

سورة البقرة (الآية ٣٢)

إهداء

إلى والدي العزيز ... إلى والدتي العزيزة أطال الله في عمرهما وأمدهما بالصحة والعافية.

إلى زوجي وابني يَحْيَى.

إلى إخوتي الأعزاء.

وإلى جميع أساتذتى وزملائى وإلى كل من مد لي يد العون... أهدي هذا العمل المتواضع.

شكر وتقدير

الحمدلله رب العالمين والصلاة والسلام على خاتم النبين والمرسلين نبينا محمد وعلى آله وصحبه اجمعين، أما بعد، فإنى أشكر الله سبحانه وتعالى على ما من به على ووفقني إلى انجاز هذه الدراسة، وأتقدم بجزيل الشكر والعرفان إلى **الأستاذ الدكتور/ مصطفى جلال مصطفى والأستاذ الدكتور/ مدحت محمد أحمد عبد العال رئيس قسم الإحصاء، الرياضة والتأمين** اللذان تفضلا بالاشراف على هذه الدراسة حيث قدما لي العناية والرعاية والنصح والإرشاد طيلة فترة إعداد الدراسة، فلهما مني جزيل الشكر والعرفان والوفاء.

كما أتقدم بالشكر والتقدير إلى **الأستاذ الدكتور/ عبدالله أحمد عبد الغالى** على تفضله بقبول مناقشة هذه الدراسة وإثرائها بالملاحظات والتوجيهات القيمة.

كما أتقدم بالشكر والتقدير إلى **الأستاذ الدكتور/ عمرو إبراهيم عبد الرحمن الأتربي وكيل الكلية لشئون خدمة المجتمع والبيئة** على تفضله بقبول مناقشة هذه الدراسة وإثرائها بالملاحظات والتوجيهات القيمة.

وختاما لايفوت الباحثة أن توجه خالص الشكر والتقدير إلى أسرتها الكريمة وكل من ساعدها وقدم لها العون في إنجاز هذة الدراسة.

الباحثة

مستخلص

منى محمود محمد عبادة، استخدام الأساليب الإحصائية لاقتراح محافظ الاستثمار في الأوراق المالية، ماجستير الإحصاء التطبيقى، جامعة عين شمس، كلية التجارة، قسم الإحصاء والرياضة والتأمين، ٢٠١٤.

تعتبر البورصة من أحد أشكال الاستثمار الشائعة نظرًا للعائد المتوقع المرتفع ولكن تقابله مخاطر مرتفعة، لذلك تقترح عدة دراسات بعض الوسائل لمساعدة المستثمر فى اتخاذ القرار. فتكوين المحفظة من أحد الوسائل المهمة لتقليل المخاطرة فى الاستثمار فى الأوراق المالية، حيث أنه يتم تقسيم أصول المستثمر وتوزيعها على مجموعة من الأسهم لتنويع المحفظة حيث أن أى خسارة ناتجة من سهم ما قد يؤدى إلى انخفاض أو معادلة السلوك التاريخى لأحد الأسهم الأخرى لذلك هناك عدة أدوات يمكن الاعتماد عليها لتحديد أداء وعلاقات الأسهم ببعضها البعض ويطلق على أحد هذه الأدوات التنقيب فى البيانات **Data Mining** والتى أثبتت فعالية وفائدة كبيرة فى تطبيقها على القطاع المالى. وتهدف هذه الدراسة إلى اقتراح عدة بدائل لمحافظ الأوراق المالية. وتقنيات التنقيب فى البيانات المستخدمة فى هذه الدراسة هى التقنيات غير الموجهه **Unsupervised** التحليل العنقودى **Cluster analysis**وشبكات **Kohonen** والتى ستساعد فى تصنيف سوق البورصة المصرية. وأيضاُ تحديد العلاقات بين القطاعات المختلفة لسوق البورصة المصرية باستخدام التنقيب فى قواعد الربط **Association Rules Mining**. وأوضحت شبكات **Kohonen** نتائج أفضل فى تصنيف سوق البورصة المصرية من التحليل العنقودى وأيضاُ تطابق بين نتائج **Kohonen** ونتائج التنقيب فى قواعد الربط.

قائمة المحتويات

قائمة الجداول..... VII

قائمة الأشكال التوضيحية VIII

قائمة المصطلحات IX

قائمة الاختصارات XII

الباب الأول:الإطار العام للبحث..... ١

١-١- مقدمة ٢

١-٢- طبيعة وأهمية البحث..... ٣

١-٣- أهداف البحث ٤

١-٤- خطة البحث وأسلوب الدراسة..... ٤

١-٥- تحديد وتوصيف مجتمع البحث..... ٥

١-٦- حدود الدراسة ٦

١-٧- مصادر البيانات ٧

١-٨- الدراسات السابقة..... ٨

الباب الثانى:التنقيب في البيانات ١١

٢-١- مقدمة ١٢

٢-٢- التنقيب فى البيانات..... ١٣

٢-٣- مصادر البيانات ١٤

٢-٤- عملية التنقيب في البيانات Data Mining Process ١٥

٢ـ ٥ـ معالجة البيانات: ١٩
٢ـ ٦ـ التقنيات المستخدمة في التنقيب في البيانات ٢٠
٢ـ ٧ـ مجالات تطبيق التنقيب فى البيانات..... ٢٣
الباب الثالث:محافظ الاستثمار في الأوراق المالية ٢٦
٣ـ١ـ مقدمة ٢٧
٣ ـ٢ـ مقدمة عن المحافظ الاستثمارية ٢٩
٣ـ٣ـ عوامل نجاح الاستثمار ٣٠
٣ـ٤ـ المحفظة الاستثمارية ٣١
٣ـ٥ـ تطوّر مفهوم المحفظة ٣٢
٣ـ٦ـ أدوات الاستثمار التي تتشكل منها المحافظ الاستثمارية ٣٤
٣ـ٧ـ أدوات الاستثمار ٣٥
٣ـ٨ـ تكوين المحافظ الاستثمارية..... ٣٦
٣ـ٩ـ تمويل المحفظة..... ٣٧
٣ـ١٠ـ المقومات الأساسية للمحفظة الشاملة المتوازن..... ٣٨
٣ـ١١ـ القائمة الاسترشادية التفصيلية لقرارات الاستثمار في المحفظة ... ٤٢
٣ـ ١٢ـ قائمة المحفظة ٤٣
الباب الرابع: تقنيات التنقيب في البيانات ٤٥
٤ـ ١ـ مقدمة ٤٦
٤ـ٢ـ التحليل العنقودي Cluster analysis ٤٧
٤ـ٣ـ المسافات والتشابه Distances and Similarities..... ٤٨

٤-٤- مراحل التحليل العنقودي: ٤٩
٤-٥- مقاييس التشابه والاختلاف Similarity& Dissimilarity ٥١
٤-٦- مصفوفة التقارب Proximity Matrix ٥١
٤-٧- الشكل البيانى للتقارب Proximity Graph..... ٥٢
٤-٨- مقاييس المسافة للبيانات الرقمية ٥٢
٤-٩- مقاييس التشابه والاختلاف بين المجموعات..... ٥٤
٤-١٠- مقاييس التشابه والاختلاف بين المتغيرات..... ٥٥
٤-١١- خوارزميات التحليل العنقودى Clustering Algorithms ٥٥
٤-١٢- تمثيل التحليل العنقودي الهيكلي..... ٥٦
٤- ١٣- الشجرة الهيكلية Dendrogram..... ٥٧
٤-١٤- خوارزميات التحليل العنقودي الهيكلي المجمع ٥٨
٤-١٥- الطرق الهيكلية المقسمة..... ٦٣
٤-١٦- تقييم خوارزميات التحليل العنقودي..... ٦٣
٤-١٧- الشبكات العصبية الاصطناعية ٦٨
٤-١٨- تصميم الشبكة العصبية ٦٩
٤-١٩- اختيار خوارزمية التعلم Learning Algorithm ٧١
٤-٢٠- تعلُم الشبكة..... ٧٤
٤-٢١- أنواع تعلم الشبكة ٧٥
٤-٢٢- الشبكات العصبية الموجهة..... ٧٥
٤-٢٣- البرسبترون Perceptron ٧٦

٤-٢٤- العنصر الخطى المتكيف Adaptive linear(Adaline) ٧٨
٤-٢٥- البيرسبترون متعدد الطبقات(MLP) ٨٠
٤-٢٦- دوال التنشيط (التحويل) غير الخطية ٨١
٤- ٢٧- تعلم شبكة MLP ٨٣
٤-٢٨- طرق تعلم شبكة MLP ٨٣
٤-٢٩- الشبكات العصبية غير الموجهة ٨٦
٤-٣٠- بناء الشبكات غير الموجهة ٨٧
٤-٣١- تعلم الشبكات غير الموجهة ٨٨
٤-٣٢- الخرائط المنظمه ذاتيًا Self Organizing Maps ٩٠
٤-٣٣- سودوكود الخوارزمية Pseudo-code of SOM ٩٤
٤-٣٤- خريطة تدفق لمراحل SOM ٩٥
٤-٣٥- التنقيب عن قواعد الربط Association Rules Mining ٩٦
٤-٣٦- قياس قواعد الربط ٩٨
٤-٣٧- خوارزمية Apriori ١٠٢
الباب الخامس:الدراسة التطبيقية لتقنيات التنقيب في البيانات . ١٠٦
٥-١- مقدمة ١٠٧
٥-٢- نتائج التحليل العنقودي ١٠٨
٥-٣- نتائج شبكة كوهونن Kohonen Network ١٢٨
٥-٤- نتائج التنقيب في قواعد الربط ١٣٣
الباب السادس ١٤٠

النتائج والتوصيات ١٤٠
أولًا: النتائج..... ١٤١
ثانيًا: التوصيات..... ١٤٣
المراجع ١٤٤
أولاً: المراجع العربية ١٤٤
ثانياً: المراجع الأجنبية ١٤٦
الملاحق..... ١٥٠

قائمة الجداول

جدول ٤-١: مؤشرات تقارب المجموعات ٦٥
جدول ٥-١: مصفوفة المسافة ١١٠
جدول ٥-٢: جدول التجميع – الربط المنفرد ١١١
جدول ٥-٣: توزيع الشركات على العناقيد – الربط المنفرد ١١٥
جدول ٥-٤: جدول التجميع – الربط الشامل ١١٦
جدول ٥-٥: توزيع الشركات على المجموعات – الربط الشامل ١١٩
جدول ٥-٦: جدول التجميع – الربط المتوسط ١٢٠
جدول ٥-٧: توزيع الشركات على العناقيد – الربط المتوسط..... ١٢٣
جدول٥-٨: جدول التجميع – WARD ١٢٤
جدول ٥-٩: توزيع الشركات على العناقيد – WARD ١٢٧
جدول ٥-١٠: توزيع الشركات على المجموعات ١٣١
جدول ٥-١١: توزيع الشركات على مجموعات (يتبع) ١٣٢

قائمة الأشكال التوضيحية

شكل ٤-١ : المهمات المختلفه للتنقيب في البيانات ٤٨
شكل ٤-٢ : مراحل التحليل العنقودي ٥٠
شكل ٤-٣: الطرق المختلفة للتحليل العنقودي ٥١
شكل ٤-٤: التحليل العنقودي المجمع والمقسم ٥٦
شكل ٤-٥: 5-TREE ٥٧
شكل٤-٦: DENDROGRAM ٥٧
شكل ٤-٧: شبكات التغذية الأمامية ٧٠
شكل ٤-٨: الشبكات المتكرره ٧٠
شكل ٤-٩: الشبكات المترابطه ٧١
شكل ٤-١٠ : دالة THRESHOLD ٧٢
شكل ٤-١١: دالة STEP ٧٢
شكل ٤-١٢: دالةPIECEWISE LINEAR ٧٣
شكل ٤-١٣: دالةLOGISTIC SIGMOID ٧٣
شكل ٤-١٤: دالةSYMMETRIC SIGMOID ٧٤
شكل ٤-١٥ :شبكة PERCEPTRON ٧٨
شكل ٤-١٦ : توضيح تعلم ADALINE ٧٩
شكل ٤-١٧ : دالة LOGISTIC SIGMOID ٨١
شكل ٤-١٨: دالة SYMMETRIC SIGMOID ٨١
شكل ٤-١٩: دالة GAUSSIAN ٨٢
شكل ٤-٢٠: دالة GUASSIAN COMPLEMENT ٨٢
شكل ٤-٢١ :شبكة KOHONEN ٩١
شكل ٤-٢٢: خريطة تدفق SOM ٩٥
شكل ٤-٢٣: خوارزمية A-PRIORI ١٠٢
شكل ٤-٢٤: خوارزمية شذب القواعد ١٠٤
شكل ٥-١: الشجرة الهيكلية – الربط المنفرد ١١٣
شكل ٥-٢: مسافات الربط عبر المراحل – الربط المنفرد ١١٤
شكل ٥-٣: الشجرة الهيكلية – الربط الشامل..... ..١٢٧
شكل ٥-٤: مسافات الربط عبر المراحل – الربط الشامل..... ...١١٨
شكل ٥-٥: الشجرة الهيكلية – الربط المتوسط..... ....١٢٩
شكل ٥-٦: مسافات الربط عبر المراحل – الربط المتوسط ١٢٢
شكل ٥-٧: الشجرة الهيكليه – WARD..... ١٢٥
شكل ٥-٨: مسافات الربط عبر المراحل - WARD..... ١٢٦
شكل ٥-٩: توزيع الأسهم على نيورونات الشبكة ١٣٠
شكل ٥-١٠: البناء الهيكلي لشبكة KOHONEN..... ١٣٠
شكل ٥-١١ :مدرج تكرارى يوضح توزيع البيانات..... ١٣١
شكل ٥-١٢: شكل ثلاثي الأبعاد لقيم الداعم ١٣٥
شكل ٥-١٣: شكل ثلاثى الأبعاد لقيم الثقة..... ١٤٦
شكل ٥-١٤: شكل القواعد ١٣٨
شكل ٥-١٥ :شبكة القواعد..... ١٤٩

قائمة المصطلحات

Activation Functions	**دوال التنشيط**
Adaptive Learning Systems	**أنظمة التعلم المُتكيفة**
Advanced Databases Systems	**نظم قواعد البيانات المتقدمة**
Agglomerative Clustering	**التحليل العنقودى المجمع**
Algorithms	**الخوارزمية**
Antecedent	**السابق**
Association Rules Mining	**التنقيب عن قواعد الربط**
Associative Network	**الشبكة المترابطة**
Attributes	**الصفات – المتغيرات**
Back propagation of errors	**الارتداد الخلفى للأخطاء**
Chained Clusters	**العناقيد المتسلسلة**
Classification	**التصنيف**
Cluster Analysis	**التحليل العنقودى**
Compact Clusters	**العناقيد المضغوطة**
Competitive Learning	**التعلم التنافسى**
Complete Linkage	**الربط الشامل**
Confidence measure	**مقياس الثقة**
Connection Weights	**أوزان التصال**
Consequent	**الاحق**
Corrective Learning	**التعلم بالتصحيح**
Cross Industry Standard Process	**عملية عبر الصناعات المعيارية**
Customer Relationship Management	**ادارة العلاقات مع العملاء**

قائمة المصطلحات (يتبع)

Dendrogram	**الشجرة الهيكلية**
Dimensionality Reduction	**تخفيض الأبعاد**
Discriminant Analysis	**تحليل التمايز**
Dissimilarity Measures	**مقاييس الاختلاف**
Divisive Clustering	**التحليل العنقودى المقسم**
Features Map	**خريطة السمات**
Fraud Detection	**كشف الغش**
Hard Clustering	**التحليل العنقودى المحكم**
Hierarchical Clustering	**التحليل العنقودى الهيكلى**
Hypertext Data	**بيانات النص التشعبى**
Input Layer	**طبقة المدخلات**
Knowledge Discovery	**اكتشاف المعرفة**
Kohonen Network	**شبكة كوهونن**
Lift Measure	**مقياس الرافع**
Memory Based Reasoning	**الاستدلال المبنى على الذاكرة**
Multilayer Perceptron	**البيرسبترون متعدد الطبقات**
Multimedia Data	**البيانات متعددة الوسائط**
Neighborhood Topology	**طوبولوجيا التجاور**
Neuron	**النيورون**
Output Layer	**طبقة المخرجات**
Partitional Clustering	**التحليل العنقودى المجزأ**
Pattern Recognition	**تمييز الأنماط**

قائمة المصطلحات (يتبع)

Perceptron	**البيرسبترون**
Proximity Matrix	**مصفوفة التقارب**
Pseudocode	**السودوكود**
Recurrent Network	**الشبكات المتكررة**
Reinforcement Learning	**التعلم بالتعزيز**
Relational Databases	**قواعد البيانات العلائقية**
Rough set Analysis	**تحليل المجموعات التقريبية**
Self-Organizing Maps	**الخرائط المنظمة ذاتيًا**
Sequential Pattern	**الأنماط المتسلسلة**
Similarities Measures	**مقاييس التشابه**
Single Linkage	**الربط المنفرد**
Steepest Descent Method	**طريقة التدرج الحاد**
Supervised Learning	**التعلم الموجه**
Support Measure	**مقياس الداعم**
Support Vector Machine	**متجه الدعم الآلى**
Transactional Databases	**قواعد بيانات العمليات**
Unsupervised Learning	**التعلم غير الموجه**
Vector Quantization	**تقسيم المتجه**
Web Click – Stream Analysis	**تحليل تدفق النقر**