ทำนายมูลค่าหุ้น Dow Jones รายสัปดาห์

( Dow Jones Index weakly prediction )

ณภัทร จารุพันธ์ 5604062630309 ตอน 2

**บทนำ**

บทความนี้ได้นำเสนอระบบทำนายมูลค่าหุ้นDow Jonesเป็นส่วนหนึ่งของวิชา Data pattern Recognition เป็นระบบที่พัฒนาขึ้นโดยใช้วิธี Regression เพื่อทำนายมูลค่าหุ้น Dow Jones ในแต่ละสัปดาห์ และใช้วิธี Trend Analysis ในการทำนายแนวโน้มของหุ้น

**1.คำนำ**

ปัจจุบันมีผู้คนมากมายนิยมลงทุนโดยการเล่นหุ้นเพราะผลตอบแทนดี และยังง่ายต่อการติดตามไม่ว่าจะอยู่ที่ไหนก็สามารถเช็คราคาหุ้นผ่านอินเตอร์เน็ตได้ตลอกเวลา แต่สำหรับผู้ลงทุนมือใหม่บางคนอาจยังไม่มีประสบการณ์ในการเลือกลงทุน ว่าควรควรลงทุนกับหุ้นตัวไหน ราคาเปิดตัวของหุ้นมีผลกับมูลค่าการซื้อขายและราคาปิดตัวอย่างไร จึงมีได้มีการจัดทำโปรแกรมการทำนายมูลค่าหุ้นเพื่อเป็นแนวทางในการประกอบการตัดสินใจในการลงทุนซื้อหุ้น

การพัฒนาระบบการทำนายมูลค่าหุ้นนั้นจะใช้ข้อมูลทางสถิติจากการเก็บรวบรวมข้อมูลในช่วงระยะเวลาหนึ่งเพื่อให้สามารถนำมาวิเคราะห์และหาสมการในการทำนายมูลค่าของหุ้นด้วยวิธีการRegression และการทำนายแนวโน้มของหุ้นด้วยวิธีการTrend Analysis

**2.ทฤษฎีและงานวิจัยที่เกี่ยวข้อง**

Regression เป็น เทคนิคที่ใช้ในการทำนายค่าตัวเลขจากข้อมูลอินพุต หลักการของ regression ก็คือการค้นหาความสัมพันธ์ (relationship) ระหว่างตัวแปรสองตัวแปร(variables) ซึ่งมักแทนด้วย x เป็นตัวแปรอิสระ(independent variable) และ y เป็นตัวแปรตาม(response variable) y = R(x) และสมการถดถอยหลายตัวแปร (Multiple Linear Regression)

Data Normalization เป็นการจัดการเพื่อให้ค่าของข้อมูลหรือคุณลักษณะมาอยู่บนบรรทัดฐานเดียวกัน ซึ่งวิธี

ที่ใช้กันอย่างแพร่หลายคือการแปลงเป็นค่ามาตรฐาน โดยคิดจากค่าเฉลี่ยและส่วนเบี่ยงเบน

มาตรฐานของข้อมูล แต่สำหรับการประมวลผลเวลาจริง (Real Time) นั้นการหาค่าเฉลี่ยและส่วน

เบี่ยงเบนมาตรฐานของข้อมูลทำได้ยาก จึงต้องหาวิธีที่เหมาะสมต่อไป

Trend Analysis เป็นเทคนิคที่ใช้วิเคราะห์แนวโน้มในการเพิ่มขึ้นหรือลดลงของข้อมูล โดยอาศัยข้อมูลทางสถิติที่เก็บรวบรวมไว้ช่วงระยะเวลาหนึ่ง ยิ่งข้อมูลมากการคำนวณแนวโน้มของข้อมูลก็จะถูกต้องมากขึ้น

**3.วิธีการดำเนินการ**

**3.1 ศึกษาและวิเคราะห์ปัญหา**

การศึกษาวิเคราะห์ข้อมูลในการพัฒนาระบบทำนายมูลค่าหุ้น ปัจจัยสำคัญที่เราสามารถทำนายมูลค่าและแนวโน้มของหุ้นได้มีความใกล้เคียงกับข้อมูลจริงนั้น ต้องมีข้อมูลที่นำมาวิเคราะห์ก่อนการทำนายมากพอสมควร ยิ่งมากยิ่งแม่นยำขึ้น เพื่อให้ได้ค่า Coefficient ที่แม่นยำมากที่สุด

การศึกษานี้เน้นไปที่การที่การ Fitting ข้อมูล 2 ข้อมูล คือ มูลค่าหุ้นเปิดตัว (open) และมูลค่าหุ้นปิดตัว (close) ที่เรียงตามเวลาแล้ว

**3.2 การรวบรวมข้อมูล**

ข้อมูลที่นำมาเป็นส่วนหนึ่งในการพัฒนาระบบทำนายมูลค่าหุ้นได้มาจาก archive.ics.uci.edu/ml/index.htmlซึ่งข้อมูลนั้นได้มีการรวบรวมและจัดเรียงอย่างครบท้วนโดยไม่มีการศูนย์หายของข้อมูลและไม่มีข้อมูลที่ผิดพลาด

ข้อมูลหุ้นทั้งหมดมีจำนวน 750 อาทิตย์ มี 2ไตรมาส ประกอบด้วยหุ้น 30 ตัวแต่ละตัวจะมีมี Attribute ทั้งหมด 16 ชนิด การวิเคราะห์จะนำข้อมูล มูลค่าหุ้นเปิดตัวต้นอาทิตย์ (open) และมูลค่าหุ้นปิดตัวปลายอาทิตย์ (close) มาใช้วิเคราะห์เป็นหลัก

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| quarter | stock | date | open | High | low | close |
| 1 | AA | 1/7/2011 | 15.82 | 16.72 | 15.78 | 16.42 |
| 1 | AA | 1/14/2011 | 16.71 | 16.71 | 15.64 | 15.97 |
| 1 | AA | 1/21/2011 | 16.19 | 16.38 | 15.6 | 15.79 |
| 1 | AA | 1/28/2011 | 15.87 | 16.63 | 15.82 | 16.13 |
| 1 | AA | 2/4/2011 | 16.18 | 17.39 | 16.18 | 17.14 |

**3.3 การวิเคราะห์ระบบ**

ข้อมูล Dow Jones Index weakly

แยกข้อมูลตามชื่อหุ้น และเรียงตามเวลา

เลือกเอา เฉพาะ Attribute Open และ Close

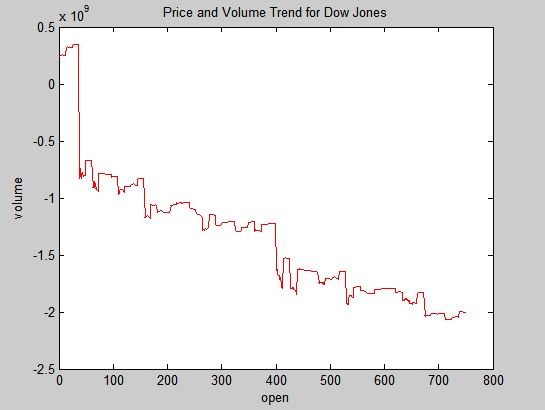
ทำการ Fittingกับข้อมูล

หาค่า coefficient

นำสมการไป

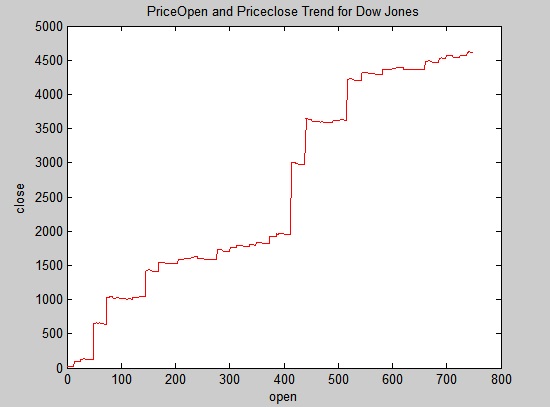
ทำนายผล

**4.ผลการดำเนินงาน**

****

**แนวโน้มราคาเปิดตัวกับมูลค่าการซื้อขาย**

แสดงให้เห็นได้ว่าราคาเปิดตัว(open)กับราคาปิดตัว(volume) ไม่มีความสัมพันธ์กัน

****

**แนวโน้มราคาเปิดตัวกับราคาปิดตัว**

แสดงให้เห็นได้ว่าราคาเปิดตัว(open)กับราคาปิดตัว(close)มีความสัมพันธ์กันมากจำเหมาะแก่การทำการวิเคราะห์แบบถดถอย(Regression analysis)เป็นอย่างมาก

**5.สรุปผล**

**Trend Analysis**

1.แนวโน้มราคาเปิดตัวกับมูลค่าการซื้อขาย ที่ราคาเปิดตัวต่ำจะมีการซื้อมากและที่ราคาเปิดตัวสูงจะมีการซื้อลดต่ำลงมาตามสัดส่วน

2.แนวโน้มราคาเปิดตัวกับราคาปิดตัว ราคาเปิดกับราคาปิดตัวจะต่างกันไม่มากช่วงเวลาหนึ่ง แต่พอเวลาผ่านไปช่วงราคาปิดตัวจะเพิ่มขึ้นแบบเอ็กซ์โพเนนเชี่ยล และพอผ่านไปอีกช่วงเวลาหนึ่งจะกลับมาค่อยๆเพิ่มขึ้นทีละน้อย แล้วจะวนซ้ำไปจนถึงช่วงเวลาที่กลับมาค่อยๆเพิ่มขึ้นที่ละนิด

จึงเหมาะสมแก่การใช้สมาการถดถอยดีกรี 3 เป็นอย่างยิ่ง

**6.เอกสารอ้างอิง**

# Data Pattern Recognition : An Overview on Time Series อาจารย์ สถิตย์ ประสมพันธ์