**基于机器学习和改进Black-Scholes期权定价模型的前沿市场股价预测**

摘要：布莱克-斯科尔斯(Black-Scholes)期权定价模型(BSOPM)长期以来一直被用于股票期权的估值，以求取股票的价格。本文利用BSOPM提出了一种比较分析方法和数值方法，求出了看涨期权和看跌期权的价格，并将这两种价格作为前沿市场股票的买入价和卖出价，对股票价格(收盘价)进行了预测。对模型进行了修改，找出了计算前沿市场股价的“执行价”和“到期时间”等参数。为了验证改进的BSOPM算法的有效性，我们采用了机器学习的方法，使用了RapidMiner软件，采用了不同的算法，如决策树算法、集成学习算法和神经网络算法。已经观察到，使用机器学习对收盘价的预测与使用BSOPM获得的预测非常相似。由于Black-Scholes-Merton方程包含了连续变化的风险和红利参数，因此机器学习方法比BSOPM方法具有更好的预测性。我们还对波动率进行了数值计算。当股票价格因定价过高而上涨时，波动性会以惊人的速度增加，当波动性变得非常高时，市场往往会下跌，这可以用我们改进的BSOPM来观察和确定。在类比量子物理薛定谔方程(和热方程)的基础上，对所提出的改进的BSOPM进行了解释。

关键词：Black-Scholes期权定价模型；Black-Scholes方程；机器学习；数据挖掘；股价预测；薛定谔方程

1.引言

20世纪初，路易斯·巴谢利耶(Louis Bcherier)首先开始了寻找代表期权定价的估值公式的工作。20世纪60年代，期权定价的研究取得了重大进展，这反映在一些金融经济学家所做的大量工作中。金融数学中最基本和最强大的工具之一是布莱克-斯科尔斯方程，它用于计算资产价格。欧式看涨期权的价格可以用Black-Scholes方程[1，2]的解来确定。此外，本文还推导了非线性Black-Scholes方程来模拟投资组合套期保值过程中的交易成本[3，4]和大额交易者的反馈效应[5-8]。期权被定义为一种合同或证券，赋予买方或持有者权利，而不是义务，以特定的价格(称为执行价)和特定的日期(称为股票到期日)买卖资产[9，10]。在期权定价中，如果股票的价格高于执行价，那么期权往往有更高的价值，因此期权被行使，但如果股票的价格低于执行价，期权的价值就会降低，因此期权肯定会到期而不被行使[9]。

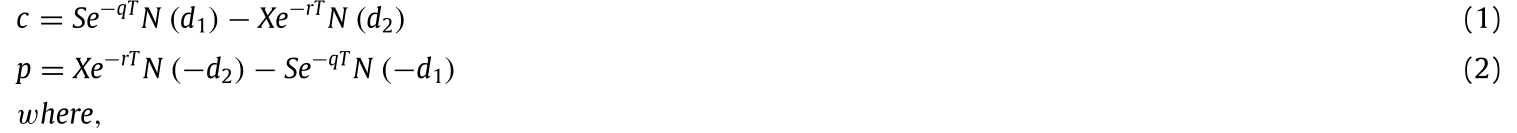
从历史上看，期权是金融工具中最复杂、最深奥的领域之一，因为它的不确定性，是一个非常令人困惑的金融学问题。然而，随着期权定价模型的发展，期权交易变得更加容易，期权定价不再被视为深奥的金融工具。期权可以有两种类型-看涨期权和看跌期权。如果执行价低于当前股价，看涨期权就被认为是实值期权。平值期权是当前股价与执行价相同或接近，而虚值期权是股价低于执行价。在看跌期权的情况下，如果股票的当前价格低于执行价，就说它是在实值。因为平值的情况下，执行价格等于或接近当前的股价，最后，如果是虚值，股票价格则高于执行价格。然而，在现实中，这并不总是发生在行使期权的情况下，现实与BSOPM的理论解释更加背离。这种背离是由供求、银行利率、投机压力等因素造成的，这些因素也影响着市场交易的发生。然而，Black-Scholes开发了一个模型，在该模型中，他们主张通过动态套期保值来消除“风险”参数，从而对期权进行估值[11]。因此，包括风险手段在内，理论价格与实际交易价格的偏差。

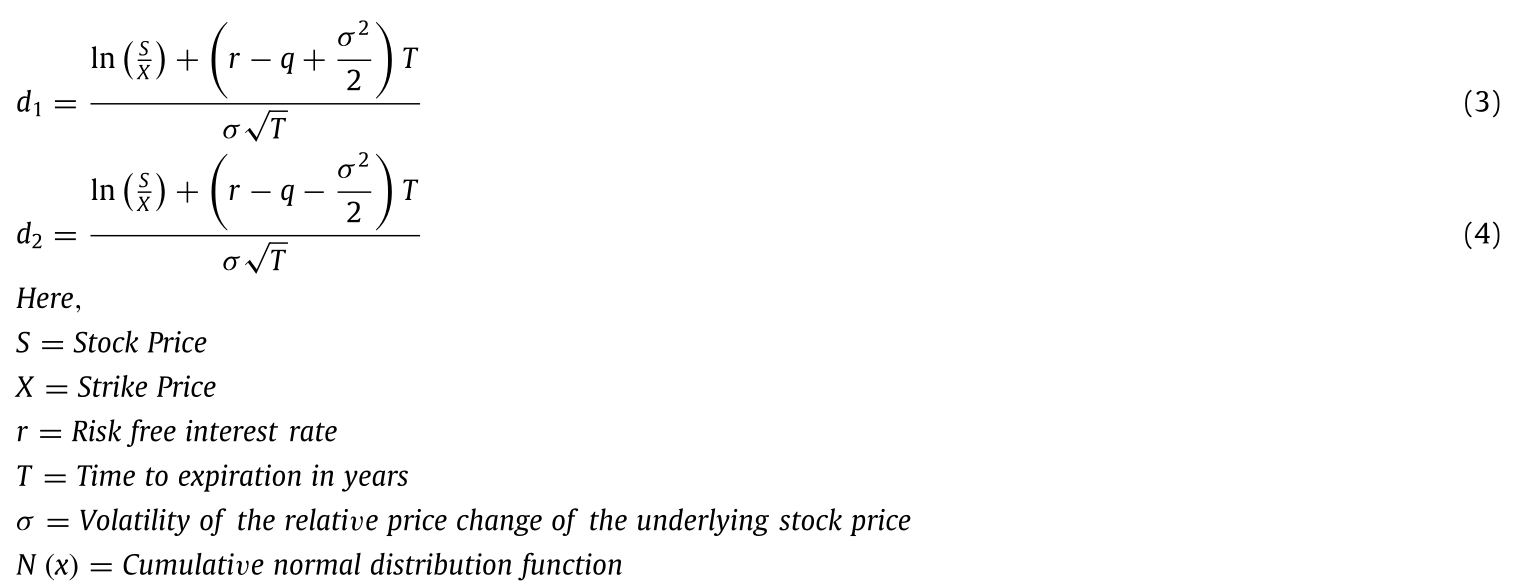
利用Black-Scholes公式对某一特定公司在前沿市场的股价进行预测的研究至今尚未见报道。原因是：前沿市场不遵循新兴市场的方法，在新兴市场进行期权交易，可以使用BSOPM预测像股票价格这样的资产价格。原因是BSOPM有执行价、到期时间和波动率等参数，这些参数在前沿市场上甚至都不存在。然而，在预测前沿市场和衍生市场的股价时，机器学习和数据挖掘扮演着非常重要的角色。近几十年来，机器学习和数据挖掘技术的应用促进了各种商业智能系统的发展。原因是：算法可以通过机器学习来训练和开发，这样它就可以几乎准确地根据历史数据预测股票的未来价格，并通过预测和描述来显示结果。另一方面，由于Black-Scholes期权定价模型在很大程度上是基于布朗运动的概念从物理学中推导出来的，人们从量子力学的角度构造和建立了许多不同的方法来获得Black-Scholes偏微分方程。

为了利用Black-Scholes期权定价模型预测前沿市场的股价，我们提出了一种方法，只需简单地改变和修改Black-Scholes方程的“执行价”、“到期日”和“波动率”参数。我们考虑了”美式期权”交易方法，期权可以在任何时间行使，因此它更适合预测股价，因为股票在前沿市场的交易是连续的，没有任何到期日。这项研究的一个有趣发现与Black-Scholes方程的“波动率”等参数有关。研究发现，股市崩盘与股票波动性的突然增加有直接关系。如果波动性持续增加，每个月都会随着达卡证券交易所等前沿市场公司股价的大幅上涨，那么可以理解，在可预见的未来，市场很可能会崩盘或下跌。因为波动性不可能长期维持在高水平，这是因为人们不是每天都以同样的方式交易股票，所以股票的价格必须有一些增值和减价，这是由于定价过高造成的。为了观察波动率的上升和下降，我们使用了MATLAB。原因是：该软件在计算和显示波动率(标准差)时考虑了某一前沿市场特定公司股票的收盘价(股价)。从MATLAB获得的结果是基于股票的收盘价和股票交易的相应日期，因此交易者更容易看到股票波动性的变化。这对交易者观察市场走势并从股票中获利非常有用。用Black-Scholes方程得到的结果与前沿市场上特定股票的原始股价进行了比较。为了验证Black-Scholes方程的结果，我们在RapidMiner中使用了数据挖掘和机器学习技术，使用决策树、神经网络、线性回归和集成学习等算法和方法对来自前沿市场不同证券交易所的11家公司的股票未来价格进行了预测。我们使用机器学习模型获得的最好的准确率是通过使用集成学习模型。最后，在类比薛定谔方程的基础上，从量子物理的角度解释了Black-Scholes方程的变化和修正。

2.Black-Scholes-Merton方程

Black和Scholes推导了期权定价公式，利用该公式可以确定期权的理论价值[9]。然而，该模型是以一定的参数为基础的，具有一定的假设条件。后来，R.C.默顿进一步扩展了股息支付和行权价格变动的模型[1]。公式如下：





然而，这个等式只正确处理欧式期权，因为它与美式期权不同，它有一个到期日。在本文中，我们将考虑美式期权定价方法，利用Black-Scholes-Merton模型的不同方法，通过改变一些参数来计算前沿市场的股票价格，这样会更好地适用于前沿市场。

3.方法论

BSOPM假设股票价格在任意有限时间间隔服从对数正态分布，因为BSOPM中的股票价格在连续时间内遵循随机游走，其中股票价格与变异率成正比[9]。BSOPM假设资产价格服从具有常数漂移和波动性的几何布朗运动。为期权等工具找到公平和无风险价格的第一个解决方案是由Bachelier在1900年提出的[12]，他首先使用布朗运动对股票价格变动进行数学建模，其中他使用中心极限定理推导出股票价格变动的正态分布[13]。因此，对于每天的股票价格[14]，我们可以写为

今日股价=昨日股价∗

其中，‘e’是指数项函数，‘u’被定义为周期性的每日收益率。它是指当天资产价值增加或减少的比率。因为资产的周期性回报率是一个随机数字，所以为了塑造走势并确定未来的股价，我们使用了一个模拟随机走势的公式。这主要是Bachelier用来描述股票价格变化的随机过程，现在我们称之为布朗运动[13]。布朗运动假设资产有两个部分，而BSOPM认为，在利率不变且已知的情况下，股票收益率的变异率保持不变。因此，为了计算基于BSOPM和布朗运动的周期性日收益率[u]，我们可以这样写：

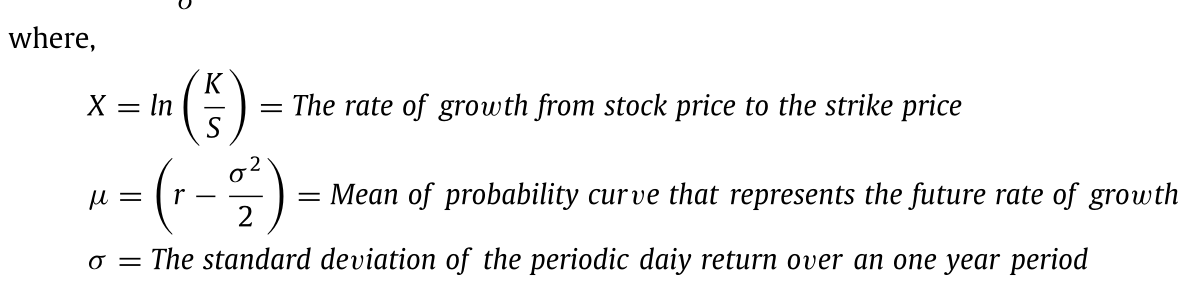


这里，第一部分是常数漂移，第二部分是随机随机分量。所以，未来的股价是，

未来股价=当前股价∗

因此，从没有无风险套利的观点来看，股票将会随着时间的推移以恒定的速度上涨，或者如果股票的风险被剔除，股票将以无风险的速度上涨。但随着前沿市场上的人们随机买卖股票，波动性也随之增加，因此计入风险回吐就意味着计入了波动性。这意味着，资产每天都可以随机增加或减少。中心极限定理指出，在一定的条件下，独立随机变量的大量迭代的平均值将是正态分布的，而不管潜在的分布是什么。每个独立随机变量都具有明确的期望值和方差。这意味着，如果从总体中抽取特定公司的大量简单随机样本，并根据每个样本计算平均值，则这些样本均值的分布将假定为正态概率分布。换句话说，如果我们做一个周期性的日收益率图，那么这个图就会形成一个正态分布的钟形图。所以我们假设未来价格的日变化率也是正态分布的。因此，未来日周期收益率的曲线是正态分布曲线，其中漂移是平均值，历史标准差是假设的未来标准差，这就是Black-Scholes公式中的布朗运动所暗示的。布朗运动是指，如果我们绘制未来周期日收益率的曲线图，我们假设，曲线图将形成一条正态分布钟形曲线，以漂移为平均值，以历史标准差为未来标准差。

未来收益率的总概率由正态分布曲线表示，因此我们可以很容易地用它来计算股票价格在到期日高于或低于期权执行价的概率。所以，我们找到了期权到期时，当前股票价格处于执行价所需的增长率，然后看看增长率落在正态分布曲线的什么位置。要做到这一点，我们要找出增长率与股价、执行价和预期增长率之间的标准差有多大。这被称为标准Z分数。概率和统计中的Z分数定义为



在图1中，曲线下的总面积表示未来增长率的总概率。Z分数右侧的百分比区域表示期权到期时股价达到或高于执行价的概率，而左侧区域表示期权到期时股价是否低于执行价的概率。

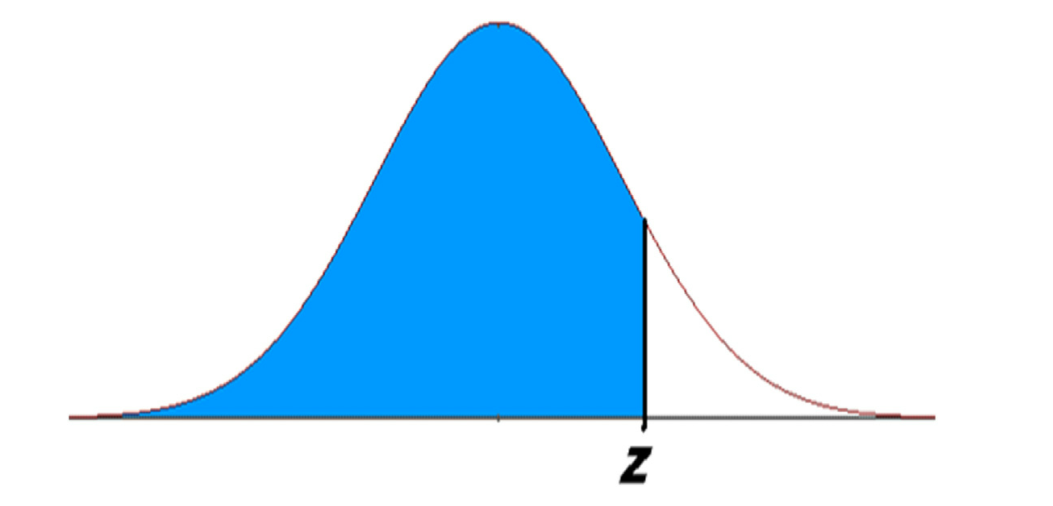


图1.概率分布曲线图

3.1找到‘’执行价‘’参数

期权所有者可以买卖标的证券的期权的固定价格就是执行价[9]。当涉及到期权估值时，执行价是最重要的因素之一，因为没有执行u；货币 价，就不可能确定期权是有价值的还是毫无价值的。前沿市场的问题是没有执行价，因为市场交易的不是期权，而是股票。因此，我们根据方法论中描述的对数正态分布，从数学上推导出了执行价格。

若,则BS公式中的执行价为

我们可以找到每日价格变化(X)，即当天股票的今天收盘价除以公司特定股票昨天收盘价的自然对数。所以, X = ln(Today′s close price/Yesterday′s close price)

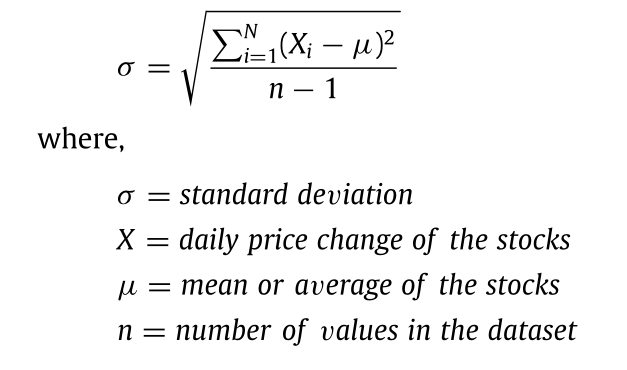
到期时间是BSOPM中最重要的关键因素之一，对BSOPM的股价有显著影响。随着时间的推移，如果股票的价格不随时间变化，期权的价格也会下降[9]。值得注意的是，时间(理论上)是以年为单位的，但期权交易员也会在剩余到期天数的情况下工作。前沿市场的问题是没有到期时间，因为它没有交易的选择权。由于股票没有到期日，人们随机地、连续地交易股票，这实际上是一个永无止境的过程。人们主要根据一定的参数来交易公司的股票，如最后交易价(LTP)、高、低、开盘价、收盘价、价值、股息、成交量、每股收益(EPS)、市盈率(P/E)、公司的长期和短期贷款、市场资本等。因此，买家可以购买某一公司的股票，并可以终身持有，可以随时以任何价格进行交易。因此，在调整到期时间时，我们考虑了全年的月度交易日。因此，它是以一个月的总交易日除以全年的总交易日来计算的。在本文中，我们考虑了使用BSOPM对未来任何一天的股票价格进行预测的方法。例如，如果买家想要在2016年3月1日购买非洲工业公司(Afric Industries SA)的股票，那么准确的日历天数2016年3月1日就会被考虑在内。让我们考虑以下情况-从2016年1月到2016年2月，卡萨布兰卡证券交易所的交易持续了42天。也就是说，2016年3月1日是第43天。从2015年3月1日开始，截至2016年2月最后一个交易日的总交易日为一年，总计约261个交易日。那么3月1日的股价时间是

Time input in BS equation =43/261= 0.164

所以，所花的时间仅仅是股票在一年的总交易日内交易的确切交易日。但是，分母边保持不变，但分子会改变。因此，通过这种方式，随着任何一天的股票价格连续改变时间，看涨期权和看跌期权的价值就被确定了。

3.3寻找“波动率”参数

在衍生品市场中起核心作用的参数是波动率，因为它直接关系到股票价格的走势。随着波动率的增大，股价的波动范围比低波动率股票的波动幅度要大得多。期权是一种非线性证券，由拥有波动性信息的交易者进行交易[15]。因此，交易者通过利用股票获利，由于波动性较大，股票的利润率较高，但这也包括资产价格的高风险。这种波动率有两种类型--一种是历史波动率，另一种是隐含波动率。Black-Scholes模型的隐含波动率是使Black-Scholes价格与市场价格匹配的波动率。也就是说，隐含波动率来自市场。历史波动率是由过去市场价格的时间序列派生出来的波动率。但不可能知道股票未来的波动率进行交易，因此交易员利用历史波动率来交易期权。前沿市场的问题是，在达卡证券交易所(DSE)、科伦坡证券交易所(CSE)、巴基斯坦证券交易所(PSX)、卡萨布兰卡证券交易所等不同的证券交易所都没有提到波动率，因此需要计算波动率，以便将其作为Black-Scholes方程的输入。波动率是标准差，可以通过两种方法确定-一种是使用直接公式，即



第二种方法是使用MATLAB SIMULINK。本文采用MATLAB SIMULINK方法，在数据统计中直接给出标准差。为了确定波动性，一个月的收盘价在Y轴上，相应的日期在X轴上。因此，MATLAB在Tools下的数据统计中显示了价格的最小值、最大值、平均值、中位数、模式、标准差和范围。这是通过输入参数”time”的相互关系来完成的。由于时间是在一年内，所以波动率是按一年的每个月分开计算的。最后，将12个月的波动率相加，作为Black-Scholes方程的输入[附录S1给出了计算股票一年期间波动率的MATLAB代码]。使用MATLAB预测看涨期权和看跌期权的公司Afric Industries SA(AFI)一年的波动率图表和对应值如表9所示。

多头头寸与投资市场有关，因为投资者对未来资产价格上涨的预期。投资者对多头头寸的关注主要集中在银行对某项资产价格的提升上。

空头头寸与多头头寸策略截然相反，因为它涉及借入证券，然后在交易者预期资产价格会以某种方式下降的情况下出售证券。因资产价格下跌而做空银行的投资者。但卖空是有风险的，因为买家必须从经纪人那里借入证券，然后通过在公开市场上出售证券来获利。然后，在证券价格降至低于买方在市场上出售时的价格后，买方必须回购证券，并再次不得不将证券卖回给经纪人。投资者在预期股价会下跌的时候做空一只股票，并可以利用下跌的股票从中获利。因此，由于卖空涉及风险，投资者更喜欢多头头寸而不是空头头寸，因为多头头寸的风险比空头头寸低得多，因为股价永远不会跌破零美元。

虽然在前沿市场上没有做多和做空这样的术语，但在现实中，这种情况确实发生了。这是因为前沿市场的交易员实际上会买入股票，买家想持有多久就持有多久。因此，买家实际上可以根据市场走势选择多头头寸和空头头寸。但事实是，买方从未在前沿市场向经纪人借过股票，因此买方没有义务在获利后返还股票。利用Black-Scholes-Merton方程，我们提出了一个确定股票买入价和卖出价的方法。因此，‘’看涨期权‘’和‘’看跌期权‘’参数，实际上是购买证券的权利和出售证券的权利，在前沿市场的情况下成为特定公司股票的买入价和卖出价。然后将这两个市场价格平均，就可以确定股票的收盘价[参见表8]。这是因为任何一家公司的股票收盘价通常是通过对收盘前最后几分钟发生的最后20个市场价格的平均来计算的。附录S3中使用Black-Scholes方程对所有11家公司的原始收盘价和预测收盘价进行了比较。

3.5股息率

在本文中，从Black-Scholes-Merton方程观察到，任何股票的价格都会随着股息率的增加而下降。当任何一家公司股票的历史收盘价与包括股息率在内的其他参数一起作为输入输入到Black-Scholes方程中，以求出该股票另一天的理论收盘价时，可以看出，由于高股息率，理论收盘价比实际收盘价有所下降。也就是说，派息股票的看涨期权和看跌期权价值往往比非派息股票的价值要低。由于我们把看涨期权和看跌期权的平均值看作是某只股票的收盘价，所以当看涨期权和看跌期权的价值下降时，理论值也会偏离原来的看涨期权和看跌期权的价值，因此我们从Black-Scholes方程得到的收盘价也会偏离原来的收盘价。所有公司股票的股息率(在Black-Scholes方程中作为输入)使用汤森路透软件收集并显示在表10中。所有公司股票的股息率是使用汤森路透基于每日频率收集的，以每股股息占股价的百分比表示。

3.6无风险利率

在现实中，没有一个术语叫做无风险利率，但是Black-Scholes-Merton偏微分方程假定短期利率是已知的，并且随着时间的推移是恒定的[9]。这就是为什么我们使用汤森路透(Thomson Reuters)软件收集所有数据，认为国库券是各自国家相应央行给出的所有公司的无风险利率，如表11所示。

4.机器学习方法

机器学习是在数据挖掘、文本挖掘、预测分析和决策系统中发挥广泛关键作用的最重要、最基本的方法之一。机器学习和数据挖掘技术可以集成到商业智能系统中，以帮助做出现实生活中的决策[16-21]。机器学习可以分为两个子类，即表示和泛化。我们已经接近泛化类别，因为它完全基于训练数据集[22]来预测看不见的数据的准确性。在本文中，我们使用了广泛使用的软件RapidMiner来处理和分析股票价格，因为它支持数据挖掘过程的所有步骤[23]。预测股市走势和格局的努力一直是一项极具挑战性的工作[24-28]。在RapidMiner软件中，数据分析通常使用图表、曲线图、图表和表格来执行，在这些图表中，用户可以很容易地将输出可视化，还可以在一个或多个属性和模型之间进行比较。但对于预测未来股票价格的机器来说，必须训练机器从给定的数据集学习，根据不同的算法创建哪些模型，从而实现预测[22]。为此，必须通过向模型提供不同的算法和要学习的训练数据集来开发模型，因此将基于数据集的概率分布函数和频率给出输出[22]。在RapidMiner这个从大数据集中获取意想不到的关系的平台中，数据挖掘是最强大的研究领域之一[29]。股票价格预测是应用金融学中最复杂的领域之一，机器学习在预测未来股票价格中起着至关重要的作用。为了最大限度地增加资本收益，最大限度地减少损失以获得最佳产出，需要准确预测股票市场价格的趋势[30]。我们在RapidMiner中使用了三种学习算法，用于创建预测模型，以预测11家公司的未来股价[参见表1]，并提供了大量输入数据集来训练这些模型[各个网站的数据参考见附录S4]。这些数据集是从四家证券交易所收集的[见表10]。我们的预测涉及六个属性[参见表1]来分析和预测未来的股票价格，其中收盘价作为“标签”属性，日期作为“ID”。



5.预测模型

5.1.决策树

决策树算法是股票价格预测中最重要的树模型之一。该运算符生成可用于分类和回归的决策树模型[31]。该运算符将属于不同类别的值分开进行分类，并且在回归的情况下，它将它们分开，以便以最佳方式减少所选参数标准的误差。通过将主数据集分成两个子集[23]，可以很容易地学习它。RapidMiner提供生成决策树模型，它自动生成树模型，以确保对我们所实验的公司的收盘价有更好的预测。为此，我们将我们的数据集分开，分别用于Excel电子表格格式的测试和培训[参见表2]。为了根据以前的数据预测任何一家公司某月的股价，我们提供了历史价格作为输入，以收盘价作为标签来训练模型。为了进行测试，我们只提供了将进行预测的月份的数据。模型展望如图3所示。该模型学习器通过形成树状的节点集合来工作，该集合旨在创建数值目标值的估计。在我们的例子中，这个目标值是股票的收盘价。对决策树的参数进行了修改，提高了预测收盘价的准确率。将Criteria参数设置为最小二乘，在该最小二乘上选择收盘价属性以分割和最小化节点中的值的平均值与真实值之间的平方距离。由于数据集的大量输入，树的最大深度从默认值20调优为25。最小增益调到0.15。

参数‘Minimum Leaf Size’和‘Minimum Size of Split’的值保持为默认值，而性能回归运算符的‘Main Critariary’已设置为’First’。Performance运算符用于统计性能评估，并提供回归任务的性能标准值列表。使用决策树模型获得的绩效向量如表3所示，适用于所有11家公司。从RapidMiner获得的图表显示了所有11家公司的原始收盘价和预测收盘价的比较，如图2(a)-(k)所示。

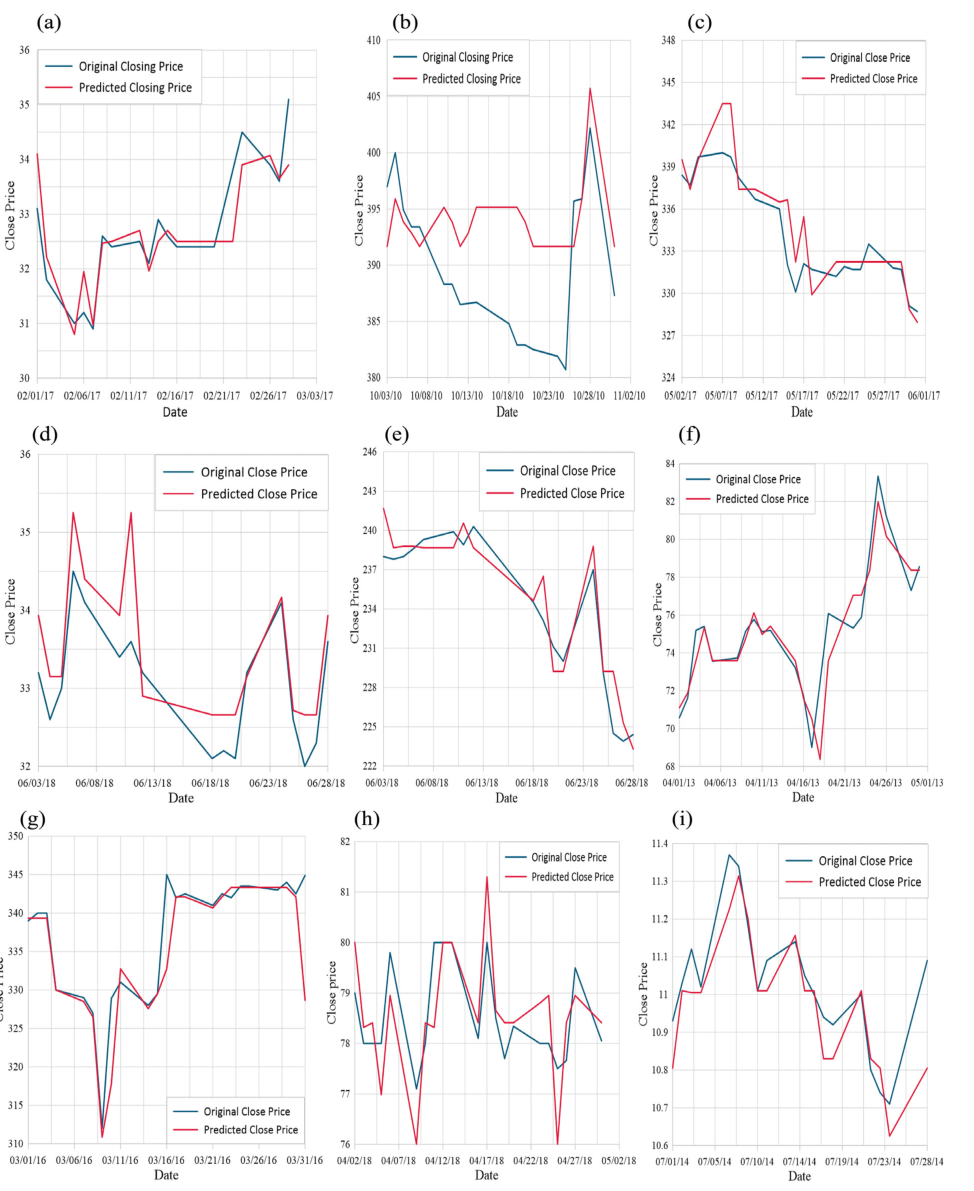
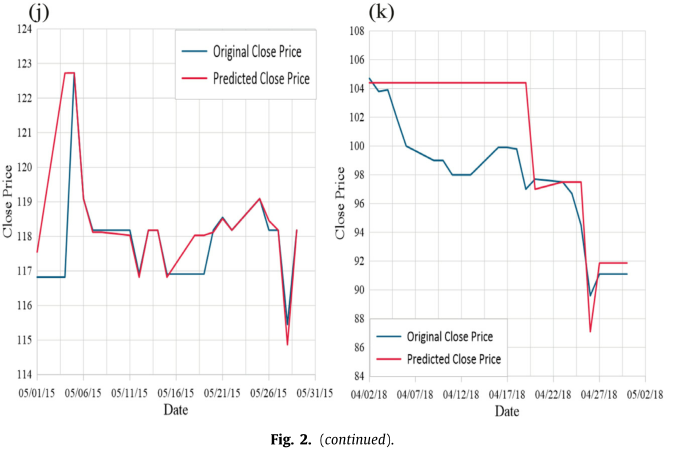
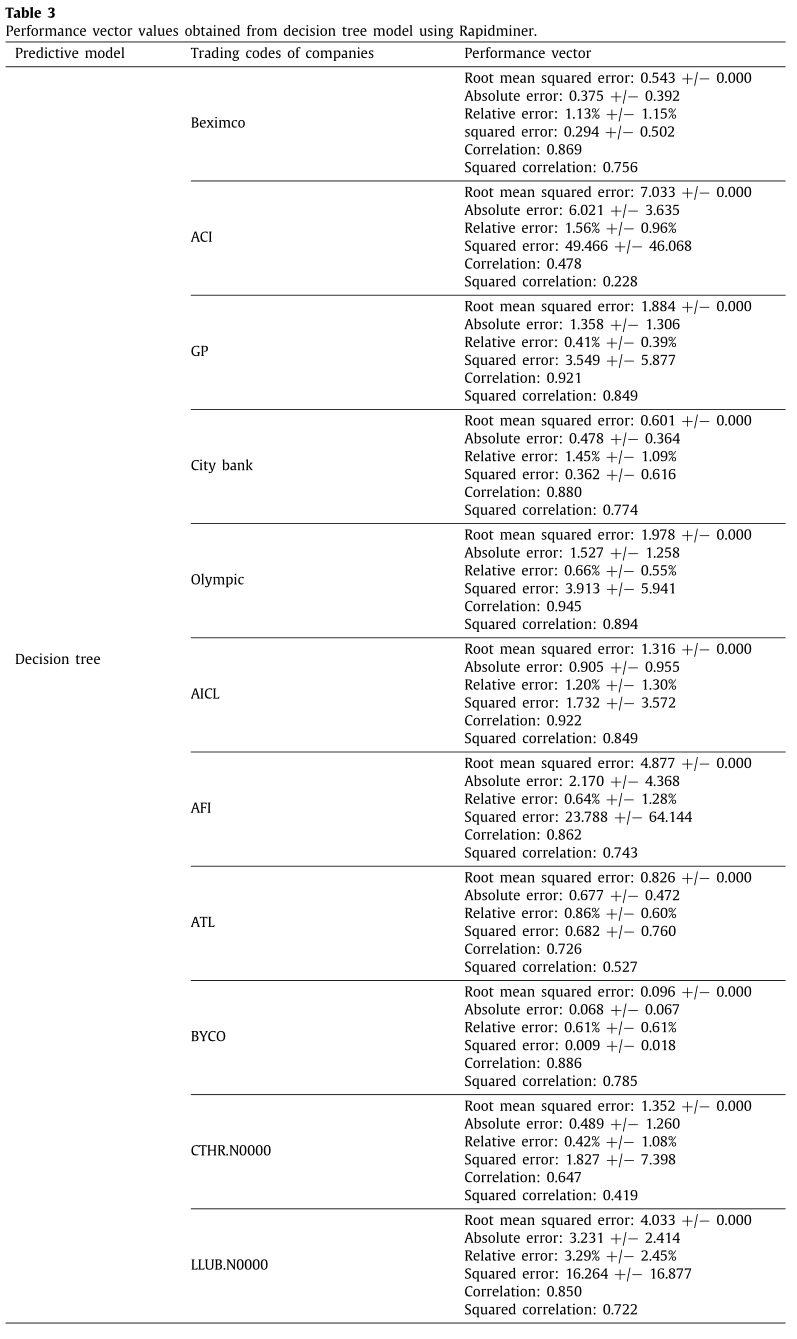


图2.使用决策树模型(A)从RapidMiner获得的Beximco(B)公司2017年2月的原始收盘价与ACI公司2010年10月的预测收盘价之间的比较。(C)GP公司2017年5月。(D)2018年6月城市银行公司。(E)2018年6月奥林匹克公司。(F)AICL公司2013年4月。(G)AFI公司2016年3月。(H)ATL公司2018年4月。(I)BYCO公司2014年7月。(J)CTHR.N0000公司2015年5月。(K)LLUB.N0000公司2018年4月。



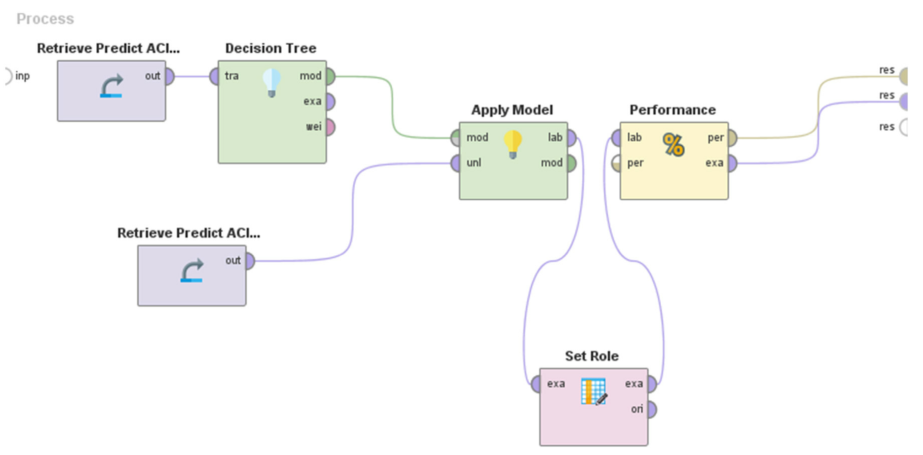


图3.用于预测公司股价的决策树模型。

5.2.神经网络

人工神经网络被定义为一个系统，它是以人脑和神经系统为模型的，因为它是一个相互连接的网络，通过各种称为神经元的节点工作。大量连接的处理单元协同工作以运行和处理输出信息。神经元的数量越少，系统的性能就越高[23]。RapidMiner提供生成神经网络模型的功能。为了在RapidMiner中执行神经网络，我们使用了与[22]中使用的模型相同的模型，如图4所示。数据集由训练和测试两个子集组成。

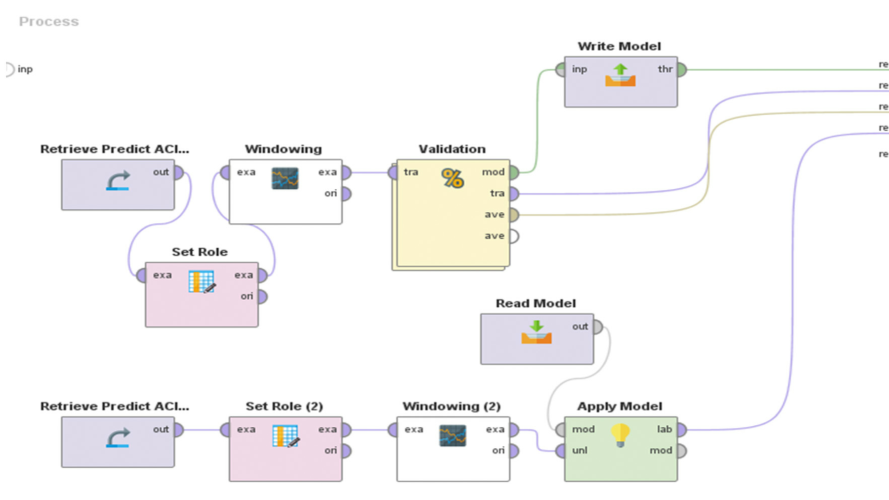


图4.用于预测公司股价的神经网络模型。

这两个数据集与表2中的决策树相同。为了通过神经网络对11家公司的股价进行预测，我们给出了所有六个属性的历史价格作为输入，对模型进行了训练。在测试集中，我们给出了要进行预测的特定月份的属性。Set Role操作符用于将属性Date的角色设置为ID，Windowing操作符用于将包含系列数据的给定示例集转换为包含单值示例的新示例集。为了相应地处理给定的输入数据集，参数序列表示、窗口大小、步长以适当的方式改变和设置。系列表示定义为系列值的编码方式，默认情况下设置为编码系列。窗口大小是使用的窗口的宽度，步长是第一个值之间的距离。这两个参数的值都设置为1。创建单个属性和创建标签已被选中，因为我们关注的是关闭属性，即，我们根据过去的关闭属性预测未来的关闭属性。Label属性设置为Close。因此，在训练窗口操作员中创建了一个标签属性，它是收盘价，而在测试端不创建标签属性，因为它将由模型预测。作为示例对时间点进行编码时，验证(滑动窗口验证)运算符用于序列预测。它使用两个示例窗口，一个窗口用于训练，另一个窗口用于测试，如图4所示。该算子的参数包括训练窗口宽度、训练窗口步长、测试窗口宽度和水平。Training Window Width是窗口中的样本数，用于训练，调整为3。Training Window Step Size是每次迭代后移动窗口的样本数，此参数调整为1。Test Width Window是用于测试的样本数，已设置为3。Horizon是从最后一次训练到第一个测试样例的增量，此值调整为1。滑动窗口验证运算符有两个阶段，即训练阶段和测试阶段。在训练阶段，学习算法采用神经网络，可以从给定的训练数据集中进行学习。用于神经网络训练的训练周期数被设置为500个周期。在学习之前，我们已经通过打乱输入数据对数据进行了排序。学习速率是每一步权重的变化量，并调整为0.3，而动量调整为0.2，这只是将上一次权重更新的一小部分添加到当前权重更新中。测试阶段包含应用模型操作符和性能参数(预测性能)。参数‘Horizon’的值已设置为1，参数‘Main Critarion’已设置为‘First’。图5(a)-(k)中显示了对比图，以直观地显示使用RapidMiner获得的原始和预测的收盘价和性能向量之间的差异和偏差。表4显示了使用RapidMiner获得的价格和性能向量。补充S2中显示了神经网络的示例。

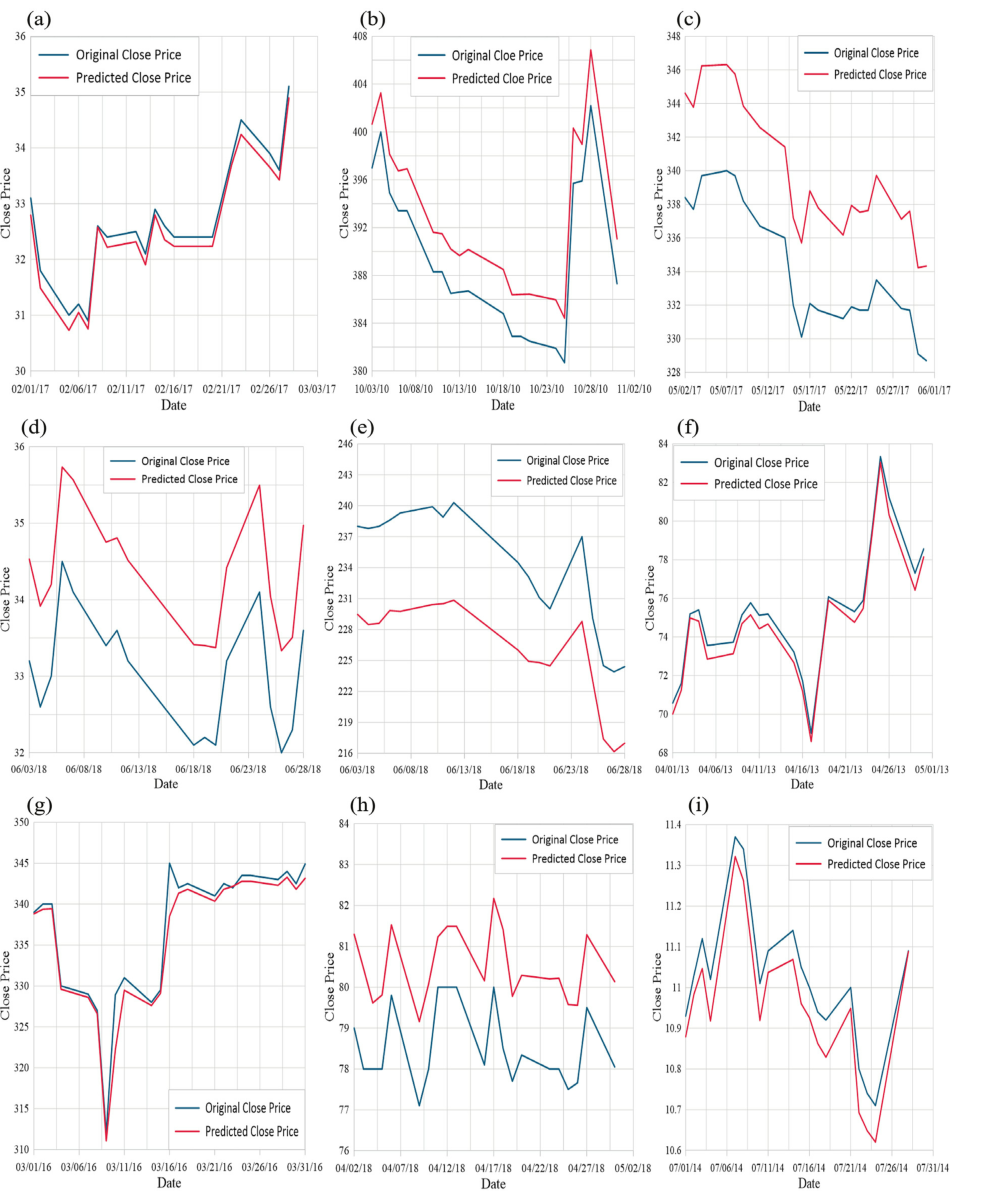
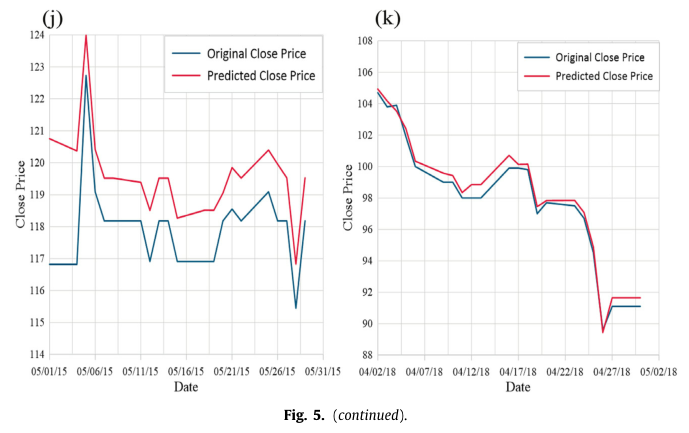
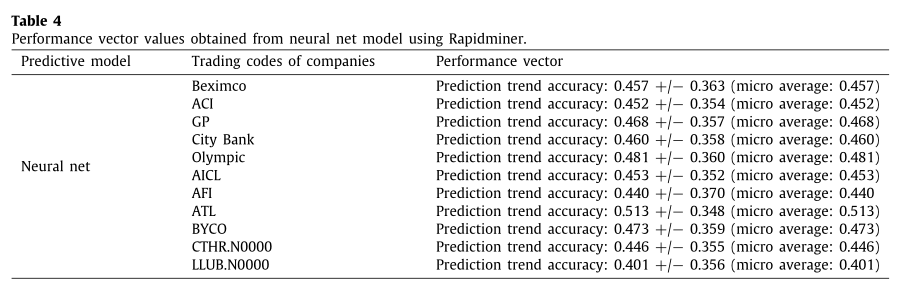
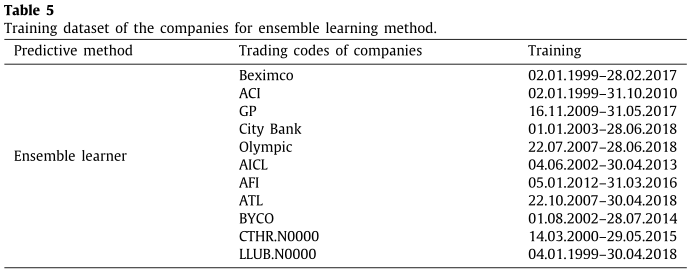


图5.使用Beximco公司2017年2月的神经网络模型(A)从RapidMiner获得的原始收盘价和预测收盘价的比较。(B)ACI公司2010年10月。(C)GP公司2017年5月。(D)2018年6月城市银行公司。(E)2018年6月奥林匹克公司。(F)AICL公司2013年4月。(G)AFI公司2016年3月。(H)ATL公司2018年4月。(I)BYCO公司2014年7月。(J)CTHR.N0000公司2015年5月。(K)LLUB.N0000公司2018年4月。





5.3.集成方法

集成学习是一种机器学习技术，其中多个机器学习算法或学习器被策略性地生成并组合在一起，以解决特定的计算智能问题。在普通的机器学习方法中，我们试图从给定的训练数据中提取一个假设，而在集成方法中，我们可以构造一组假设并将它们组合在一起。集成学习提高了模型的性能，或者它降低了不幸地选择了一个糟糕的模型进行预测的可能性。因此集成学习方法保证了最优的预测分析和最优的输出。集成方法包含多个学习器或分类器，称为基本学习器，集成方法的泛化能力比基本学习器强得多。基学习器通常是通过基学习算法从训练数据中产生的，基学习算法可以是决策树、神经网络、相对回归算法或其他类型的机器学习算法。大多数集成方法只使用一个基础学习者来产生同质学习者，但也有一些方法，它使用多个基础学习者来产生一个异质学习者，可以做出非常准确的预测[32]。集成学习降低了选择特别糟糕的算法或分类器的风险，当问题处理大数据集时，这种方法真的很有用。有许多集成方法，它们已经足够有效，可以产生准确的预测，如增压、装袋和堆积。在本文中，我们使用集成方法从给定的学习算法中获得最优输出。在机器学习和数据挖掘中，人们认为，一组模型比单个模型更好，这是集成学习的主要原则。它有助于克服个别模型的偏差和错误率，这是通过将一些弱小的学习者组合在一起产生一个强大的学习者来实现的。与集成学习的情况一样，训练数据集对模型中的错误起到了最有效的作用，因此我们取了一个大的训练集，以及六个预定义的属性和收盘价作为标签属性。该模型如图6所示。分割数据操作符产生给定示例集的所需数目的子集。示例集根据指定的相对大小划分为子集。因此，为了拆分给定的数据集，该操作符的分区参数被调优为具有两个分区，其比率分别为0.6(用于训练)和0.4(用于测试)。采样类型设置为线性采样。投票操作符是集成操作符，它也是嵌套操作符，这意味着它有一个子流程。这个子过程有三个学习器，分别是决策树、神经网络和相对回归。构建这些分类器的训练数据集如表5所示。

模型中的参数已更改。在决策树的情况下，参数是准则、最大深度。Criteria的作用是选择将根据其选择要拆分的属性的标准。我们将此值设置为最小二乘法，在该最小二乘法中选择要拆分的属性，以最小化节点中的值平均值与真实值之间的平方距离。用最大深度来限制决策树的深度，并将其调整为20。图7显示了树形图的样本叶子。对于神经网络，有500个训练周期，学习率调整为0.3，动量调整为0.2。相对回归模型是一种回归模型，它对于允许对具有大趋势的数据集进行时间序列预测非常有用，因为它学习了一种回归模型，用于相对于另一个属性值进行预测。在我们的例子中，属性值被设置为收盘价。相对回归包含线性回归运算符，其根据输入数据集计算线性回归模型，该线性回归模型被提供用于训练。然而，研究发现，通过在投票子过程中使用多元相对回归模型，提高了模型整体性能的准确性。该集合的总体预测将是单个基分类器预测的大部分。因此，为了获得基本学习者的多数票，在建立模型时设置了应用模型运算符和性能运算符，这将决定模型的性能(准确性)。性能矢量如表6所示。最后，对比图[参见。图8(a)-(k)]显示了股票原始收盘价和预测收盘价之间的比较，而在图9(a)-(k)中，给出了所有机器学习模型预测和Black-Scholes预测与原始收盘价的对比图。

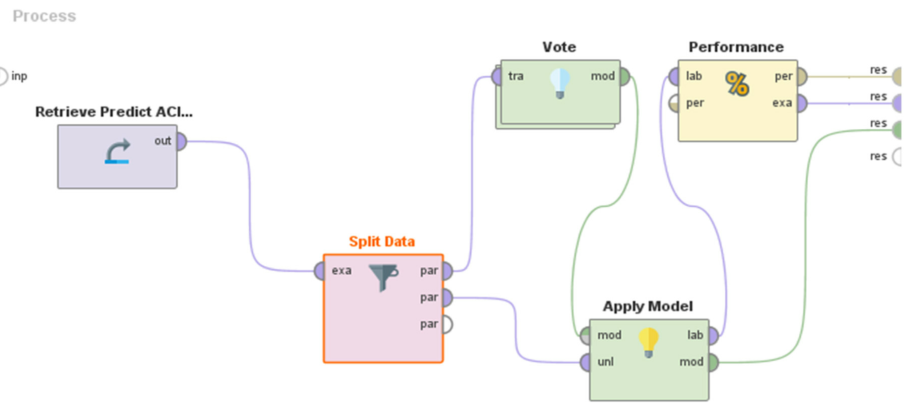
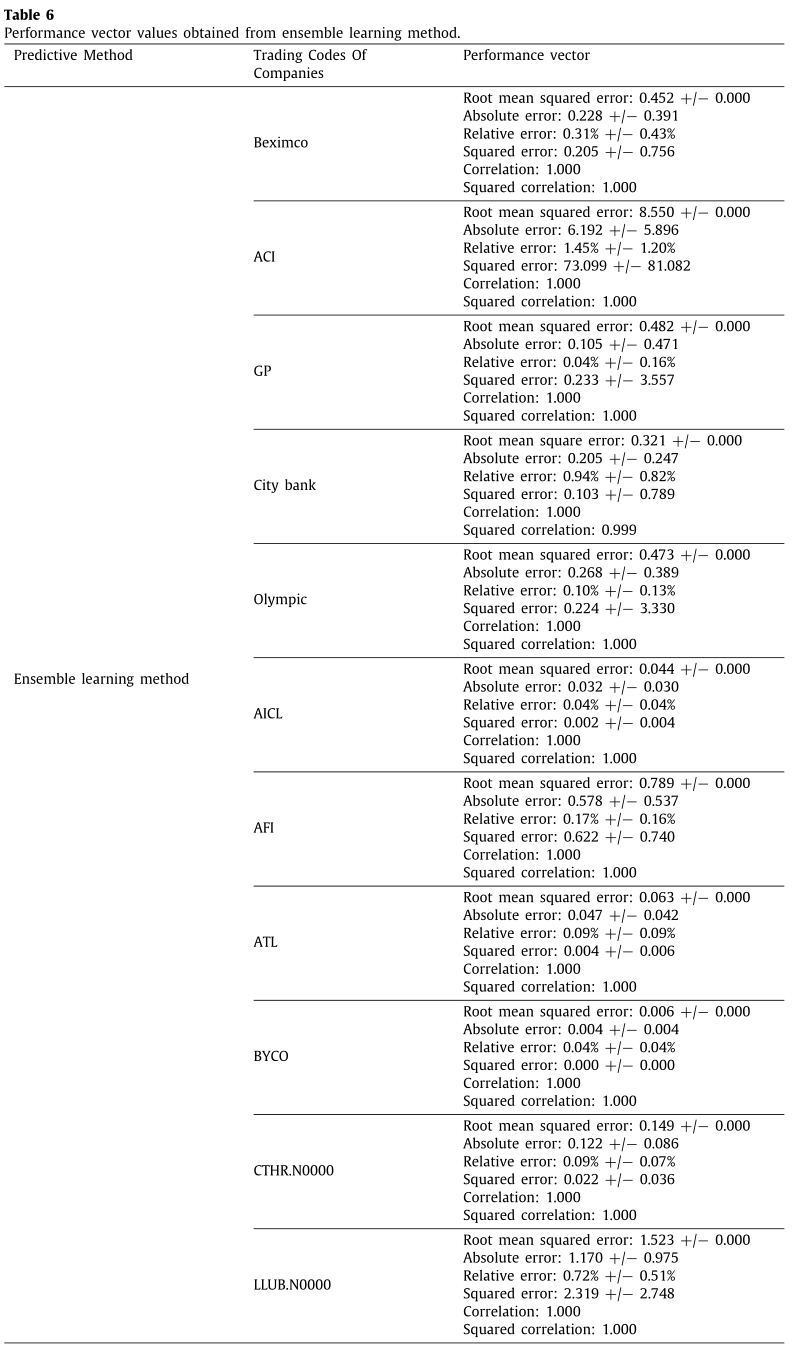


图6.用于预测RapidMiner收盘价的集成学习模型

6.从量子物理的角度解释

BSOPM的修正，可以用不同的方法来推导Black和Scholes偏微分方程，其中也包括那些利用量子物理[33-38]和其他物理学分支的方法。物理学家一直在试图使用物理思想和概念更好地理解金融动态[39-43]。在这一部分中，我们主要想关注的是‘到期时间’和‘执行价’参数变化背后的逻辑原因，以及这些变化的原因。期权价格取决于股票价格，而股票价格是一个随时间演变的随机变量[33，44]。Black，Scholes和Merton提出的期权定价方法是由随机微分方程组成的，但这是在伊藤微积分的解释范围内的。布莱克和斯科尔斯偏微分方程可以写成以下形式的扩散方程：





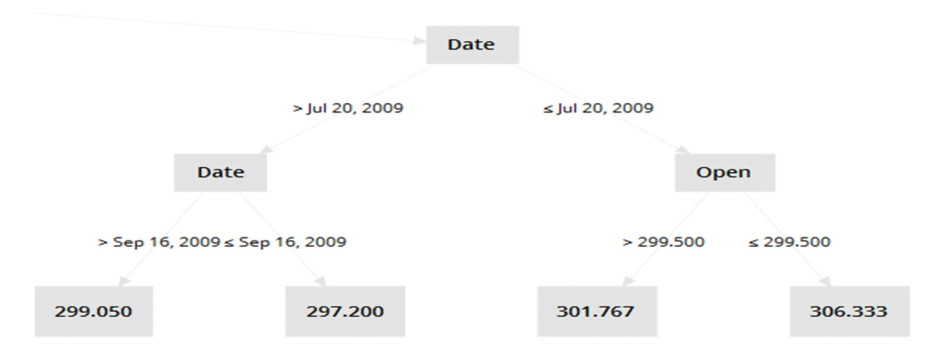
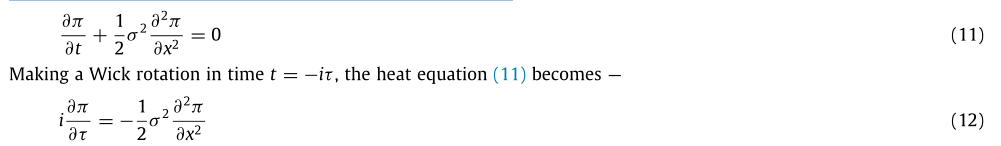


图7.集成学习方法中使用的决策树

Black-Scholes方程公式背后的基本思想是标准布朗运动过程，它由下式给出



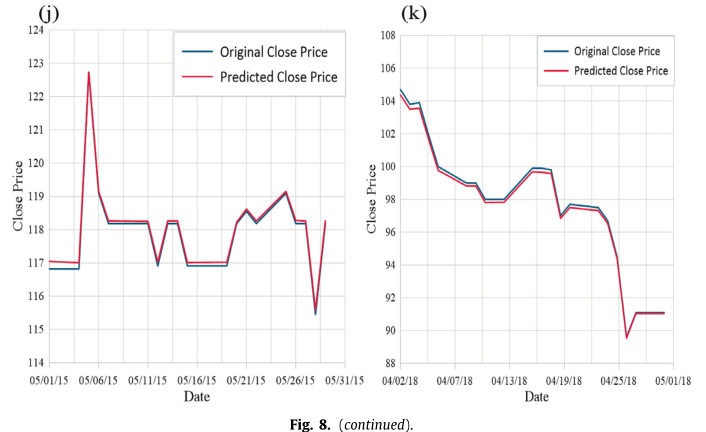
在(10)中，是一个维纳过程或标准布朗运动过程，它实际上是一个连续的随机过程。在本文中，我们在方程(5)中也有一个随机随机分量。用于查找对应股价的执行价。此外，随机方程中的随机项(9)必须是增量相关的，因为市场瞬间包含了关于未来市场演变的任何信息，这是有效的市场假说[33，44-46]。这意味着，理论价格是由白噪声[33]推动的，白噪声是白散粒噪声和维纳过程的组合[47]。BSOPM已经扩展到引入白散粒噪声[48，49]。然而，在离散时间内，白噪声的样本是具有零均值和有限方差的不相关随机变量序列，并且样本是独立的，具有相同的概率分布。但我们知道，如果每个样本的正态分布的均值为零，那么信号就是高斯白噪声。然而，在这项研究中，我们发现均值并不总是零[参见表7]，而且略大于零，这就得出结论，白噪声在BSOPM中可能不存在的可能性是有限的。此外，由于样本在时间上是连续的，Black-Scholes-Merton方程中的股票价格在时间上也是连续的。因此，在本研究中，参数“时间”是按顺序变化的，即到期时间也随股票价格的变化而按顺序变化。因此，到期时间取决于特定日期的股票价格，并相应地随特定日期的股票价格而变化。隐含波动率不仅随着时间的推移而变化，而且在任何时间点上都不同于相关期权，即一组特征仅在执行价[50]上不同的期权。第三，用Stratonovich微积分和Ito微积分得到的Black-Scholes方程是相同的，这意味着当考虑不同的方法时，时间是不变的[44]。它只随某一天的股价而变化。然而，商品或资产的未来价格是指一个人可以同意在未来的特定时间买入或卖出它的价格[51]。由于未来的价格是时间的函数，因此资产的买卖将根据时间而定，因此时间取决于特定日期的股票价格。因此，时间简单地说就是股票将在某一天交易的时间或日期。

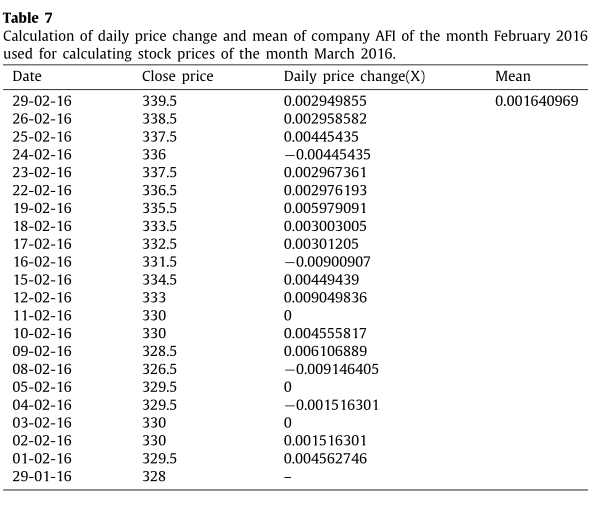
函数π=π(s，t)表示期权价格作为时间和资产价格的函数。改变坐标布莱克-斯科尔斯模型用以下方式描绘[34]- 

公式(12)是具有ℏ=1和=1/σ2的自由粒子的薛定谔方程[34]。如果我们将此方程与薛定谔方程进行比较，则可以清楚地看到方程(12)的左侧包含虚数i和Wickrotated‘Time’τ，它们也是虚数。也就是说，函数π=π(τ，x)，随着虚构的“时间”τ而变化，它代表着一个期权，它将在未来一定的给定时间内交易。这是随着股价的变化而改变时间的正当理由。同样，如果我们将这个方程与热方程进行比较，描述给定位置x处的期权的函数(π=π(τ，x)将随着时间的推移而变化，因为热将在整个空间中传播。因此，计算股票价格的时间由方程(12)右侧位置x处的未知函数π(τ，x)组成。因此，由于前沿市场的资产价格并没有固定的时间，所以我们只须花时间确定股票的交易天数。



图8.使用集成学习方法(A)从RapidMiner获得的Beximco公司2017年2月的原始收盘价和预测收盘价的比较。(B)ACI公司2010年10月。(C)GP公司2017年5月。(D)2018年6月城市银行公司。(E)2018年6月奥林匹克公司。(F)AICL公司2013年4月。(G)AFI公司2016年3月。(H)ATL公司2018年4月。(I)BYCO公司2014年7月。(J)CTHR.N0000公司2015年5月。(K)LLUB.N0000公司2018年4月。





7.结论

我们提出了一种利用Black-Scholes方程来预测前沿市场公司任何股票的价格，包括任何股票的买入和卖出价格的方法。这种方法对于找出某只股票的执行价、波动性和到期时间非常有用，因为这些参数在不交易期权的前沿市场中并不真正存在。根据任何一个交易月的股价，可以预测其他月份的股价，也可以很容易地知道趋势。要做到这一点，Black-Scholes方程所需的参数也需要相应地改变，如本文所示。取任何一年中某个月份的历史股价(收盘价)，并按时间顺序变化，就可以预测未来几个月的股价。然而，波动性和无风险利率将保持不变，并且在预测一个月的所有股票价格的整个时间内都是已知的。值得注意的是，本文用Black-Scholes方程解释的这个过程实际上是一个连续的过程。因为，根据过去某个月份的股价，我们得到另一个月的预测未来股价，然后根据该月的预测，可以实现对以后几个月的股价的预测。对于在前沿市场进行交易的交易员来说，这种方法非常有用。波动率的增加和减少对人们从股票中获得最大利润是有用的，因为它表明了市场的趋势，即公司的股票价格何时会上升或下降。最大的利润可以通过这样的想法来实现，即波动性越大，股票的价格就越高。我们已经实现的Black-Scholes预测显示，价格与当天的开盘价相似，而最初的收盘价与Black-Scholes预测的价格有偏差，因为每天都会发生连续的交易。因此，股票价格(收盘价)也在不断变化。在RapidMiner中通过数据挖掘获得的预测股价的图表与使用Black-Scholes方程获得的图表非常相似。我们使用机器学习模型获得的最好的准确率是通过使用集成学习模型。机器学习方法和Black-Scholes方法的市场走势似乎也非常相似，这导致了Black-Scholes方程对于计算和预测前沿市场公司的股价非常有用。

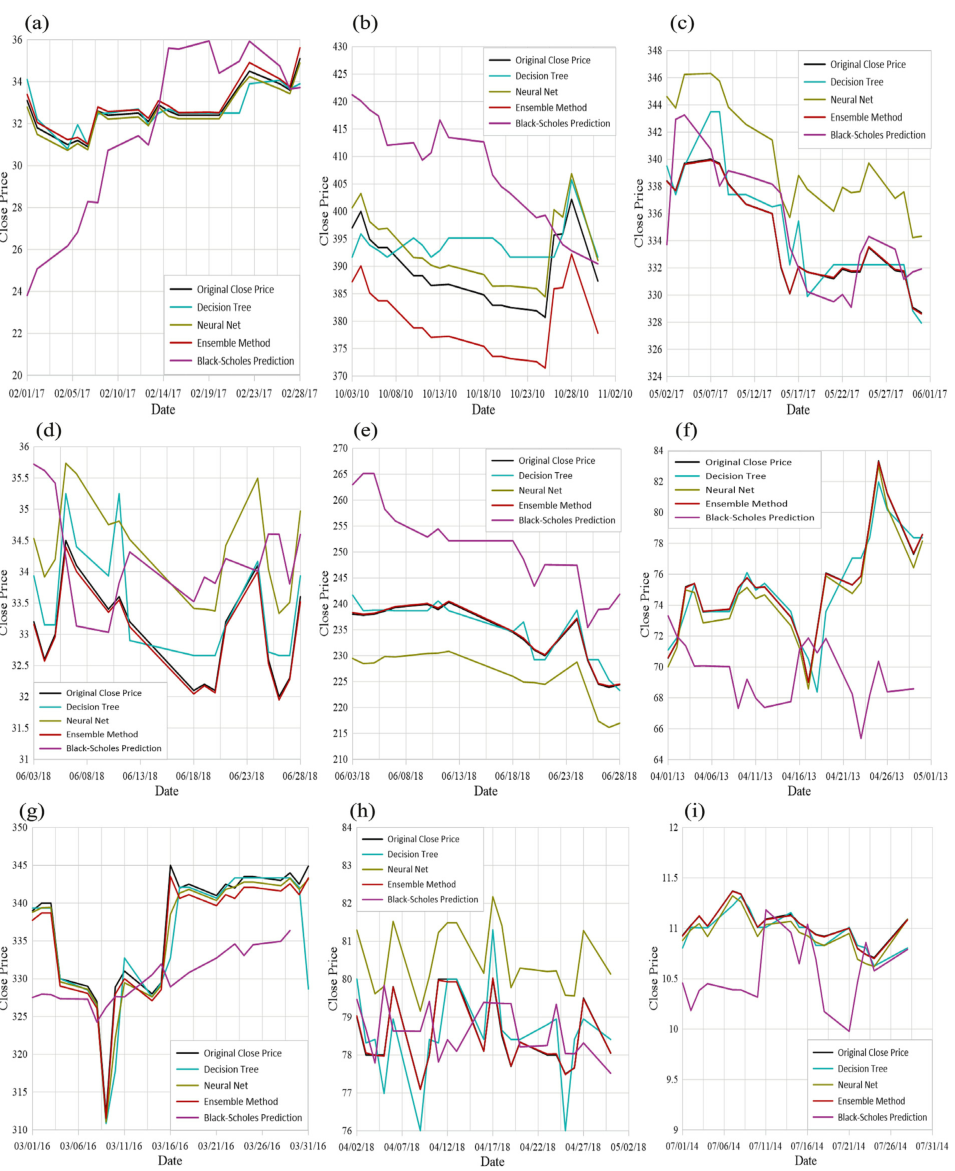
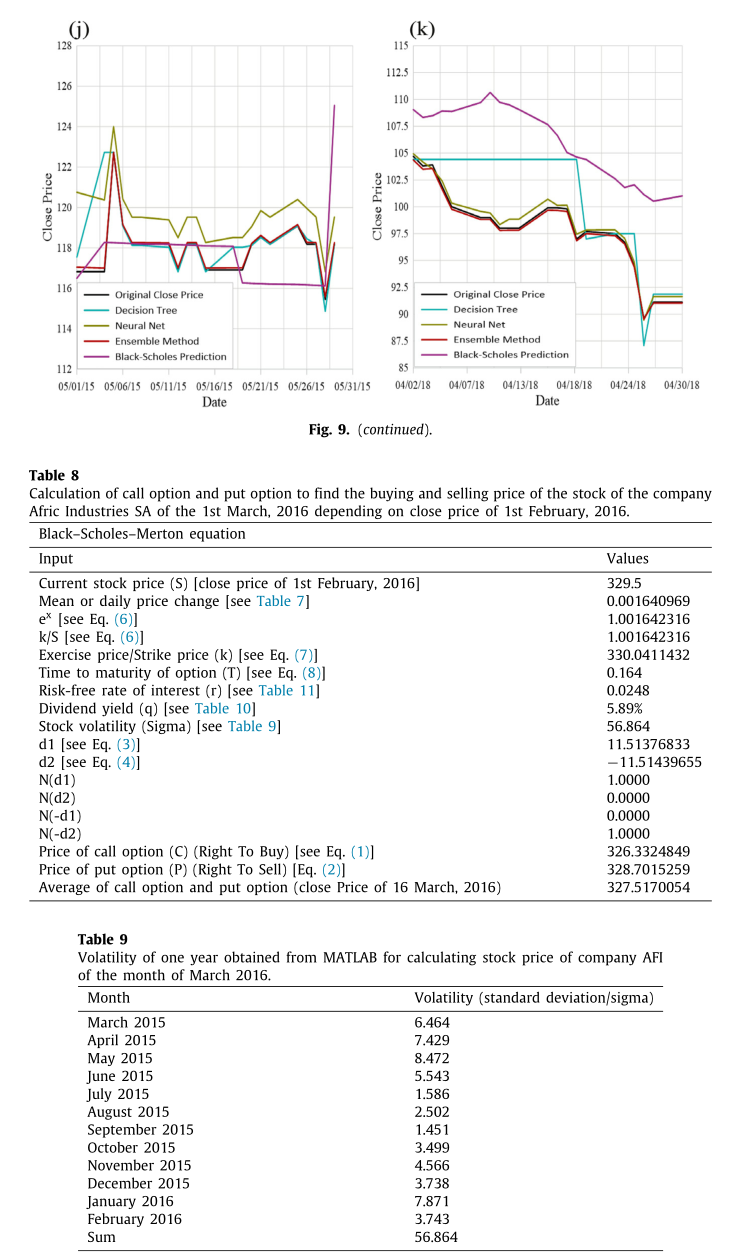
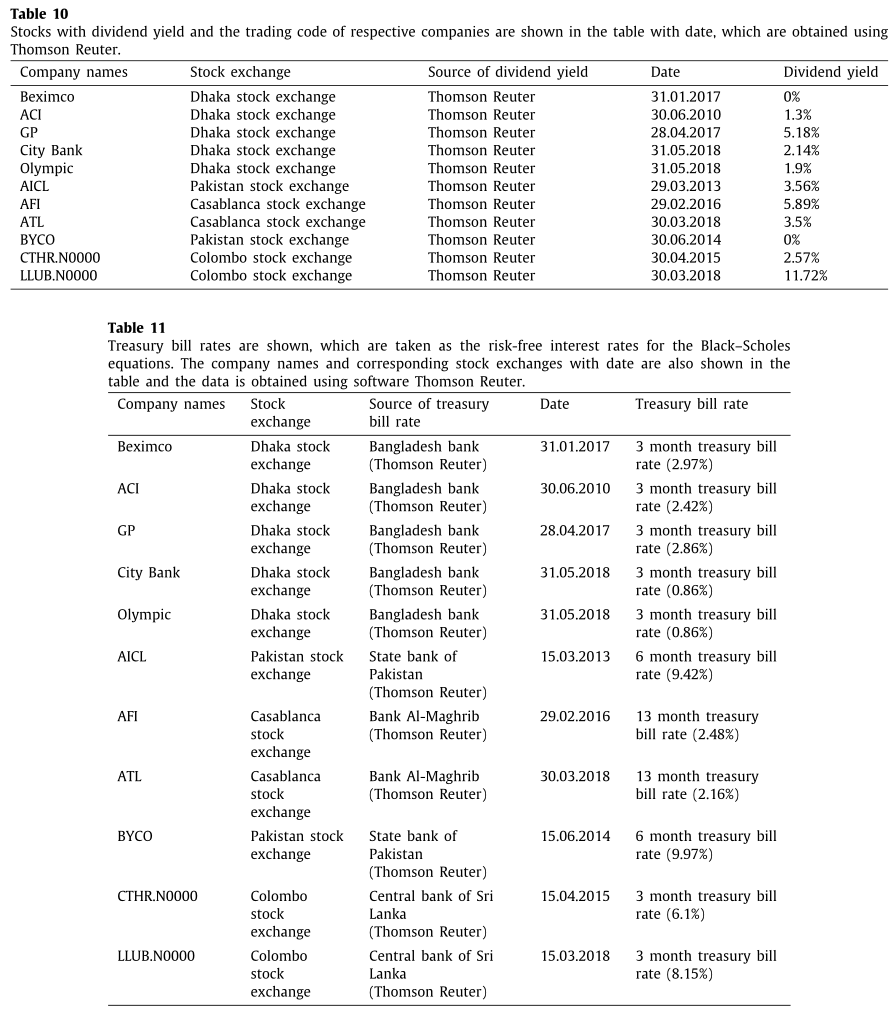


图9.Beximco公司BSOPM和RapidMiner(A)2017年2月原始收盘价与预测收盘价的比较。(B)2010年10月ACI公司的成员。(C)2017年5月GP公司的成员。(D)2018年6月公司城市银行。(E)2018年6月奥林匹克公司。(F)2013年4月AICL公司的成员。(G)2016年3月AFI公司的名称。(H)2018年4月的ATL公司。(I)2014年7月号的BYCO公司。(J)CTHR.N0000公司2015年5月的名称。(K)公司2018年4月的LLUB.N0000。





竞争利益声明作者声明，他们没有已知的竞争经济利益或个人关系，这些利益或个人关系似乎会影响本文报告的工作。

附录A.补充数据与本文相关的补充材料可在网上找到，网址是：<https://doi.org/10.1016/j.physa.2020.124444>.

References

[1] R. Merton, Theory of rational option pricing, Bell J. Econ. Manage. Sci. 4 (1) (1973) 141.

[2] B. Bensaid, J. Lesne, H. Pagès, J. Scheinkman, Derivative asset pricing with transaction costs, Math. Finance 2 (2) (1992) 63–86.

[3] P. Boyle, T. Vorst, Option replication in discrete time with transaction costs, J. Finance 47 (1) (1992) 271–293.

[4] M. Davis, V. Panas, T. Zariphopoulou, European option pricing with transaction costs, SIAM J. Control Optim. 31 (2) (1993) 470–493.

[5] R. Frey, Pefect option hedging for a large trader, Finance Stoch. 2 (1998) 115–141.

[6] G. Genotte, H. Leland, Market liquidity, hedging and crashes, Amer. Econ. Rev. 80 (1990) 999–1020.

[7] R. Jarrow, Market manipulation, bubbles, corners, and short squeezes, J. Financ. Quant. Anal. 27 (3) (1992) 311–336.

[8] E. Platen, M. Schweizer, On feedback effects from hedging derivatives, Math. Finance 8 (1) (1998) 67–84.

[9] F. Black, M. Scholes, The pricing of options and corporate liabilities, J. Polit. Econ. 81 (3) (1973) 637–654.

[10] J. Cohen, F. Black, M. Scholes, The valuation of option contracts and a test of market efficiency, J. Finance 27 (2) (1972) 399–417.

[11] Espen Haug, Nassim Nicholas Taleb, Why we have never used the Black–Scholes–Merton option pricing formula, 2007.

[12] L. Bachelier, The Random Character of Stock Market Prices, MIT Press, Cambridge, MA, 1964, p. 17,

[13] E. Sullivan, T. Weithers, Louis Bachelier: The father of modern option pricing theory, J. Econ. Educ. 22 (2) (1991) 165.

[14] InformedTrades, Black Scholes: A simple explanation [Video File], 2013, December 3, Retrieved from.

[15] S. Ni, J. Pan, A. Poteshman, Volatility information trading in the option market, J. Finance 63 (3) (2008) 1059–1091.

[16] C.Kai-Sang Leung, R. Kyle Mackinnon, Y. Wang, A machine learning approach for stock price prediction, in: IDEAS ’14 Proceedings of the 18thInternational Database Engineering & Applications Symposium, Porto, ACM, New York, NY, USA, 2014, pp. 274–277.

[17] A. Cuzzocrea, C. Leung, R. MacKinnon, Mining constrained frequent itemsets from distributed uncertain data, Future Gener. Comput. Syst. 37(2014) 117–126.

[18] C.K.-S. Leung, F. Jiang, Y. Hayduk, A landmark-model based system for mining frequent patterns from uncertain data streams, in: Proc. IDEAS,ACM, 2011, pp. 249–250.

[19] C.K.-S. Leung, F. Jiang, L. Sun, Y. Wang, A constrained frequent pattern mining system for handling aggregate constraints, in: Proc. IDEAS, ACM,2012, pp. 14–23.

[20] C.K.-S. Leung, S.K. Tanbeer, B.P. Budhia, L.C. Zacharias, Mining probabilistic datasets vertically, in: Proc. IDEAS, ACM, 2012, pp. 199–204.

[21] S. Tanbeer, C. Leung, J. Cameron, Interactive mining of strong friends from social networks and its applications in E-commerce, J. Org. Comput.Electron. Commer. 24 (2–3) (2014) 157–173.

[22] T. k, M. Wadhawa, Analysis and comparison study of data mining algorithms using rapid miner, Int. J. Comput. Sci. Eng. Appl. 6 (1) (2016)9–21.

[23] J. Han, J.C. Rodriguez, M. Beheshti, Diabetes data analysis and prediction model discovery using rapidminer, in: 2008 Second InternationalConference on Future Generation Communication and Networking, Hainan Island, 2008, pp. 96–99.

[24] I. López-Yáñez, L. Sheremetov, C. Yáñez Márquez, A novel associative model for time series data mining, Pattern Recognit. Lett. 41 (2014)23–33.

[25] W. Huang, Y. Nakamori, S. Wang, Forecasting stock market movement direction with support vector machine, Comput. Oper. Res. 32 (10)(2005) 2513–2522.

[26] M. Mostafa, Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait, Expert Syst. Appl. 37 (9) (2010)6302–6309.

[27] J. Patel, S. Shah, P. Thakkar, K. Kotecha, Predicting stock and stock price index movement using Trend Deterministic Data Preparation andmachine learning techniques, Expert Syst. Appl. 42 (1) (2015) 259–268.

[28] M. Ballings, D. Van den Poel, N. Hespeels, R. Gryp, Evaluating multiple classifiers for stock price direction prediction, Expert Syst. Appl. 42 (20)(2015) 7046–7056.

[29] B. Narayanan, M. Govindarajan, Prediction of stock market using ensemble model, Int. J. Comput. Appl. 128 (1) (2015) 18–21.

[30] L. Khaidem, S. Saha, S.R. Dey, Predicting the direction of stock market prices using random forest, 2016, CoRR, abs/1605.00003.

[31] J. Karasek, R. Burget, J. Masek, O. Benda, Genetic programming based classifier in Viola–Jones RapidMiner image mining extension, in: 201336th International Conference on Telecommunications and Signal Processing, TSP, Rome, 2013, pp. 872–876, http://dx.doi.org/10.1109/TSP.2013.6614064.

[32] Z. Zhou, Ensemble Methods, Taylor & Francis, Boca Raton, FL, 2012.

[33] V. Ivancevic, Adaptive-wave alternative for the Black–Scholes option pricing model, Cogn. Comput. 2 (1) (2010) 17–30.

[34] M. Contreras, R. Pellicer, M. Villena, A. Ruiz, A quantum model of option pricing: When Black–Scholes meets Schrödinger and its semi-classicallimit, Physica A 389 (23) (2010) 5447–5459.

[35] W. Segal, I. Segal, The Black–Scholes pricing formula in the quantum context, Proc. Natl. Acad. Sci. 95 (7) (1998) 4072–4075.

[36] E. Haven, A Black–Scholes Schrödinger option price: ‘bit’ versus ‘qubit’, Physica A 324 (1–2) (2003) 201–206.

[37] B. Baaquie, C. Corianò, M. Srikant, Hamiltonian and potentials in derivative pricing models: exact results and lattice simulations, Physica A 334(3–4) (2004) 531–557.

[38] O. Vukovic, On the interconnectedness of Schrodinger and Black–Scholes equation, J. Appl. Math. Phys. 03 (09) (2015) 1108–1113.

[39] J.P. Bouchaud, M. Potters, Theorie des risques financiers, Alea-Saclay, Paris, 1997.

[40] B.B. Mandelbrot, Fractals and Scaling in Finance, Springer, Berlin, 1997.

[41] W. Arthur, Complexity and the economy, Science 284 (5411) (1999) 107–109.

[42] T. Lux, M. Marchesi, Scaling and criticality in a stochastic multi-agent model of a financial market, Nature 397 (6719) (1999) 498–500.

[43] P. Bak, M. Paczuski, M. Shubik, Price variations in a stock market with many agents, Physica A 246 (3–4) (1997) 430–453.

[44] J. Perelló, J. Porrà, M. Montero, J. Masoliver, Black–Scholes option pricing within Itô and Stratonovich conventions, Physica A 278 (1–2) (2000)260–274.

[45] E. Fama, The behavior of stock-market prices, J. Bus. 38 (1) (1965) 34.

[46] M. Jensen, Some anomalous evidence regarding market efficiency, J. Financ. Econ. 6 (2–3) (1978) 95–101.

[47] I.I. Gihman, A.V. Skorohod, Stochastic Differential Equations, Springer, New York, 1972.

[48] R.C. Merton, J. Financ. Econ. 3 (1976) 125–144.

[49] J.C. Cox, S.A. Ross, J. Financ. Econ. 3 (1976) 145–166.

[50] M. Modisett, J. Powell, Black–Scholes option pricing model modified to admit a miniscule drift can reproduce the volatility smile, Appl. Math.03 (06) (2012) 597–605.

[51] F. Black, The pricing of commodity contracts, J. Financ. Econ. 3 (1–2) (1976) 167–179.