摘要：

在经济发展的过程中，股票市场是我们不可忽视的一部分。因为股票市场具有复杂且剧烈波动的特性，所以能帮助投资者规避风险的关于股价的预测和研究是很有必要的。众所周知，由于传统的时间序列模型ARIMA无法描述序列的非线性，其预测股价的能力有天生的缺陷；而神经网络模型具有很强的非线性泛化的能力，因此我们将利用该类模型这一优势并使用一种基于注意力机制的CNN-LSTM模型来预测股票价格。

本文构建的模型将传统的时间序列模型、基于注意力机制的卷积神经网络和长短期神经网络融合在一起，我们将它称为A-AttCL模型。该模型能充分挖掘股票市场在多个时期的历史信息，提高了预测精度。具体做法如下：我们首先用ARIMA模型对股票数据进行预处理。然后，再将其输入到基于seq2seq结构的深度学习架构中：该模型首先利用卷积提取原始股票数据的深层特征，然后利用长短期记忆网络挖掘长期时间序列特征。结果表明，混合模型更有效并且预测精度较高，可以帮助投资者或机构进行决策，达到扩大收益和规避风险的目的。

1. 绪论

股票市场在经济发展的过程中起着举足轻重的作用。由于股票具有高回报的特点，股票市场吸引了越来越多的投资机构和投资者的关注。与此同时，由于股票也具有复杂波动性的特征，它给投资人和机构带来巨大损失的惨剧也会时常发生。为了应对股票市场的巨大风险，关于股价变化的预测和研究是非常有必要的。

传统的时间序列模型ARIMA无法描述序列的非线性结构，并且还需要满足很多的前提条件，导致它在股价预测中无法取得理想的效果。近年来，随着人工智能的高速发展，越来越多的研究人员将人工智能的方法用到金融市场中。关于自然语言序列、蛋白质序列、股价序列等序列建模问题是人工智能研究的重要领域，其中神经网络是最具有代表性的一种模型，它具有很强的非线性泛化能力。

我们通常采用循环神经网络RNN对序列数据进行分析，其中长短期神经网络模型LSTM是最常用的RNN。LSTM在RNN中引入了门机制，这可以看作是对人类记忆的模拟，就像人类会记住有用的信息并且忘记无用的信息。注意力机制可以被看作是对人类注意力的模拟，正如人们会对有用的信息印象深刻，而往往会护士无用的信息。基于注意力机制的卷积神经网络（ACNN）被广泛的用在序列建模上。我们将ACNN和LSTM相结合生成一种基于自注意力机制的SEQ3SEQ模型，用它对序列数据进行编码和解码。该模型可以解决LSTM中的长程依赖问题，因此它可以更好的对长序列进行建模。LSTM可以捕捉适合它本身结构的特定的长程相关性，而ACNN可以同事捕捉到当前和全局的相关性。因此，该模型更加的灵活和可靠。

#===================

本文提出了一种混合深度学习模型来预测股价。不同于传统的混合预测模型，该模型集成了时间序列模型ARIMA与神经网络于一体，结合了两类普通模型的优点，提高了预测精度。股票数据首先经过ARIMA进行预处理，然后再输入到深度学习架构中。该训练模型是基于序列到序列框架的基于注意的CNN-LSTM模型，其中基于注意的CNN是编码器，双向LSTM是解码器。该模型首先利用卷积提取原始股票数据的深度特征，然后利用卷积提取原始股票数据的深度特征长短期记忆网络挖掘长时间序列特征。

结果表明，该模型更有效，预测精度较高，可以帮助投资者或机构进行决策，达到扩大收益和规避风险的目的。

#=====================

二、

1. ARIMA

经典的股票预测方法是基于ARMA（）模型和ARIMA（）模型。当序列是非平稳的时，ARIMA（p、q、d）对序列采用差分。

1. Deep L

在基本的前馈神经网络（FFNN）中，在t时刻的输入数据直接决定了在该时刻的输出结果，因此该模型对时间序列数据的分析能力偏弱。

在循环神经网络（RNN）中，。。。。。。结果表明，随着误差在时间维度上的传播，RNN可能会消失梯度，从而导致长程依赖问题。

正如人类可以有选择地记住信息。LSTM（通过门控激活功能，长短期记忆）模型可以选择性地记忆更新后的信息，忘记积累的信息。

序列到序列（seq2seq）模型采用自动编码器（即编码器-解码器架构）来分析序列数据。序列-序列模型（seq2seq）是按照编码-解码器结构构建的，增强了LSTM通过噪声数据学习隐藏信息的能力。在seq2seq中，编码器是一个LSTM，它将输入编码到上下文中（通常是最后hN时的隐藏状态），然后解码解码器中的上下文。在解码器中，前一时刻的输出在下一时刻被输入。

双向LSTM：为了优化解码序列，在seq2seq模型中采用了波束搜索[17]。隐马尔可夫模型（HMM）中的波束搜索和维特比算法都是基于动态规划的。根据观察和先前状态求解当前状态的最优估计，在HMM中称为解码或推理。分别求解p（xk|y1：k）和p（xk|y1：N），相当于HMM中的前向-后向算法。通过xk的分布，可以得到最优的双向估计。这就是提出了正向和反向相结合的双向LSTM的概率观点。

1. 注意力机制

正如人类通常会关注显著的信息，注意机制是一种基于人类认知系统的深度学习技术。

1. 方法
2. 预处理：

本文提出了一种混合深度学习模型来预测股票价格。与传统的混合预测模型不同，该模型将时间序列模型ARIMA与神经网络整合在非线性关系中，结合了两种普通模型的优点，提高了预测精度。

我们首先将股票数据通过ARIMA进行预处理。经过ARIMA（p=2，q=0，d=1）预处理后，可以输出一个更有效地描述状态的新序列，再将其放入神经网络（NN）中。

表I和表II中的ADF结果表明，原始序列是非平稳的，一阶差分序列是平稳的。在确定d=1后，我们需要确定ARIMA中的AR (p)和MA (q)。我们使用了自相关图（ACF）和部分自相关图（PACF）。原序列和一阶差分序列的ACF和PACF如图4和图5所示。图4显示，当阶数=2阶时，PACF截断，这意味着我们应该采用AR (2)。ACF对于任何顺序都是有长尾的，这意味着我们应该采用MA (0)。因此，p=2，q=0。

1. 训练：

训练模型是基于序列到序列框架的基于注意的CNN-LSTM模型，其中基于注意的CNN为编码器，双向LSTM为解码器。该模型首先利用卷积法提取原始库存数据的深度特征，然后利用长短期记忆网络挖掘长期时间序列特征。

Seq2seq通过编码解码器架构来抑制噪声的影响。在深度学习的基础上，该模型更有效地描述了状态的隐藏信息。LSTM从基于注意力的CNN（ACNN）编码器接收上下文信息。

ACNN编码器块由自注意层和CNN组成。经过自注意层后Q、K、V通过等式(9)计算，H通过等式(12)计算。这是LSTM解码器块的输入。编码解码层描述当前序列与之前序列之间的关系，以及当前序列与嵌入之间的关系。编码器仍采用多头机构。当第k个嵌入被解码时，只能看到第k−1个和之前的解码。这种多头机制掩盖了多头的注意力。

基于注意力的CNN（ACNN）可以捕获LSTM可能无法捕获的全局和局部依赖性，这增强了稳定性。在我们提出的编码器-解码器框架中，我们采用ACNN-LSTM结构。在人类的认知系统中，注意力通常出现在记忆之前。ACNN能够捕获长期依赖的原因，是它集成了多头自我注意和卷积。结合LSTM和ACNN可以增强结构优势和时间序列建模的能力。ACNN定位器集成了多头注意力和多尺度卷积核，可以捕获LSTM可能无法捕获的显著性，而LSTM可以更好地描述时间序列特性。

三、实验

（一）模型改进

在经过ARIMA的预处理之后，神经网络的输入是一个有一段时间间隔的二维数据矩阵，其大小为时间窗口×特征。在股票预测的实证研究中，特征包括基本的股票市场数据（开盘价格、收盘价、最高价格、最低价格、交易量、交易金额）。ARIMA处理序列和残差序列也被视为特征。

**我们采用了时间序列预测的回顾方法，回顾数字是20。这意味着时间窗口的宽度为20。LSTM的层数为5，大小为64。**

该模型通过Adam优化器进行训练，学习率为0.01。

数据从2007年1月1日至2022年3月31日的数据中选择，一天内的数据表示序列中的一个点。列车组和测试集于2021年6月22日进行划分，如图7所示。这意味着有3500个训练样本和180个测试样本。批量大小为32。

1. 预测表现

ARIMA模型的股价预测结果如图8所示。曲线和残余密度图如图9所示。

ARIMA+SingleLSTM的原序列和残差序列的损失曲线如图12和图13所示。

ARIMA+SingleLSTM模型和ARIMA+BiLSTM模型的股价预测结果如图14和图15所示。

我们所提出的模型的损失曲线如图16所示。该模型的股票价格预测结果如图17所示。

（三）

评价指标为平均绝对误差（MAE）、均方根误差（RMSE）、平均绝对百分比误差（MAPE）和r2。

这里¯Xt表示Xt的平均值。误差越小，r2越高，表示性能越好。首先，我们对不同的预训练模型和微调模型进行了比较。表三表明，我们提出的模型优于其他基线。然后，我们与现有的方法进行了比较。比较方法包括：

表四表明，所提出的神经网络的性能优于目前的方法。

1. 结论

股票市场在金融和经济发展中具有重要地位。由于股市的波动性复杂，对股价走势的预测，可以保证投资者的回报。传统的时间序列模型ARIMA不能描述股票预测中的非线性。

由于神经网络具有较强的非线性建模能力，本文提出了一种基于市场关注的CNN-LSTM和XGBoost混合模型来预测股票价格。本文模型将ARIMA模型、与注意机制的卷积神经网络、长短期记忆网络和XGBoost回归模型形成非线性关系，提高了预测精度。该模型可以捕捉到多个时期的股票市场信息。库存数据首先通过ARIMA进行预处理。

然后，采用在预训练-微调框架中形成的深度学习体系结构。预训练模型是基于序列到序列框架的基于注意的CNNLSTM模型。该模型首先采用基于注意力的多尺度卷积法提取原始库存数据的深度特征，然后利用长短期记忆网络对时间序列特征进行挖掘。最后，采用XGBoost模型进行微调。结果表明，该混合模型更有效，有助于投资者或机构实现扩大回报、规避风险的目的。