正态性检验是判断数据总体是否服从正态分布的一种检验方法，它对做统计工作之前有着重要意义，采用偏度-峰度检验法检验数据的正态分布性。

图片包含 游戏机

描述已自动生成手机屏幕的截图

描述已自动生成

有表中数据可知，CNN处理后的数据峰度和偏度更加接近于0，可以证明CNN处理后的数据相比于元数据更好的服从正态分布。

|  |  |  |
| --- | --- | --- |
|  | 峰度 | 偏度 |
| 原始数据 | 5.340550192030676 | -1.0892404692253963 |
| CNN处理后的数据 | 0.37508950888516956 | -0.6558730006217957 |

时间序列的平稳性对于有关时间序列的预测有着重要意义，非平稳的时间序列数据来预测未来的销售数据十分困难，所以将非平稳的时间序列转换为平稳的时间对预测研究有着重大意义。

对原始数据和CNN处理后的数据进行比较发现，原始数据的自相关系数在很长的延迟时期里一直保持为整数，且数值较大；CNN处理后的数据自相关系数数值较小，并且序列保持在零轴附近波动，证明了CNN处理后的数据具有较强的平稳性和相关图特征。

社交网站的手机截图

描述已自动生成手机屏幕截图

描述已自动生成

原始数据和CNN处理后的数据进行单方跟检验对比发现，原始数据检验统计量大于1%和5%的置信区间临界值，而CNN处理后的数据原始数据检验统计量远大于1%置信区间的临界值，且p值约等于0，因此可以非常好地拒绝该假设，可以很好的证明CNN处理后的数据具有一定的平稳性，对未来的预测有帮助。

*原始序列的单位*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 数据 | 检验统计量 | P值 | 1%置信区间 | 5%置信区间 | 10%置信区间 |
| 原始数据 | *-2.60611* | *0.09174* | *-3.44316* | *-2.86719* | *-2.56977* |
| CNN处理后的数据 | *-6.56076* | *8.39281 e-09* | *-3.44384* | *-2.86749* | *-2.56994* |

*根检验结果为： (检验统计量：-2.6061137887226886, P值：0.0917402438751505, 5, 513, {'1%': -3.443161545965353, '5%': -2.8671904981615706, '10%': -2.5697795041589244}, -806.059950376462)*

*CNN处理后的序列的单位根检验结果为： (-6.560762770862999, P值：8.392815975608413e-09, 12, 487, {'1%': -3.443849184997939, '5%': -2.8674931065091105, '10%': -2.569940776113236}, -1931.3273931080626)*

迪基-福勒检验（Dickey-Fuller test）

原始序列的白噪声检验结果为： (array([40.8319569]), array([1.65897903e-10]))

CNN处理后的序列序列的白噪声检验结果为： (array([116.24171911]), array([4.20757112e-27]))

CNN-LSTM正态性检测：

Kurtosis of normal distribution: 0.37508950888516956

Skewness of normal distribution: -0.6558730006217957

原始数据正态性检测：

Kurtosis of normal distribution: 5.340550192030676

Skewness of normal distribution: -1.0892404692253963

The results in different scales are shown in Fig. 2. Fig. 2 illustrates the

three-level decomposition using D-Meyer. We can see from Fig. 2(a)

that the detailed scale mainly contains the trend component, Fig. 2(b–d)

represent most of the weekly periodic components and stochastic

components. Time series forecasting can be produced by forecasting on

the trend and seasonal components separately. We thus propose that in

applying the wavelet decomposition for the sake of forecasting, it is not

necessary to decompose a given time series into too many components.

We will address this issue more in the next section.

It indicates

that there is a smaller deviation between the actual and predicted

values using the proposed model. Thus, for forecasting sales of NB,

the proposed MARS

2

VR model not only provides better forecasting

result than the four competing models in terms of forecasting

error, but also exhibits the capability of identifying important

predictor variables which may provide valuable information for

further sales management and related decisions of NB.

手机屏幕截图

描述已自动生成