计算机科学
COMPUTER SCIENCE

基于多特征融合的GRU-LSTM大学生就业动态预测

张剑，张烨

引用本文

张剑，张烨. 基于多特征融合的GRU-LSTM大学生就业动态预测[J]. 计算机科学, 2023, 50(6A): 220500056-6.

ZHANG Jian, ZHANG Ye. College Students Employment Dynamic Prediction of Multi-feature Fusion Based on GRU-LSTM [J]. Computer Science, 2023, 50(6A): 220500056-6.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.220500056

基于多特征融合的 GRU-LSTM 大学生就业动态预测

张 剑 张 烨
西安科技大学通信与信息工程学院 西安 710054

摘 要 针对高校就业预测系统大多采用单一传统特征建模而导致出现就业预测效果不佳、就业精准服务不强等问题,提出一种融合多特征因素的 GRU-LSTM 组合就业预测方法。首先,在传统预测模型特征的选择上加入了学生行为特征,并构建了多信息融合的特征向量;然后,结合不同影响因素对高校就业的贡献不同,提出了一种基于皮尔逊相关系数的多信息融合的就业预测最优特征提取方法,优化了特征子集;最后,综合考虑预测精度和预测时间两个方面的因素,提出了一种基于门控循环单元(GRU)与长短期记忆网络(LSTM)的组合预测模型 GRU-LSTM,结合 LSTM 预测精度高与 GRU 预测时间短的优点对就业数据进行高效精准预测。实验结果表明,该方法与传统方法相比,就业预测的精确率提高了 4.2%,对提高大学生就业提供了可靠的数据支撑。

关键词:深度学习;LSTM;就业预测;数据挖掘
中图法分类号 TP181

College Students Employment Dynamic Prediction of Multi-feature Fusion Based on GRU-LSTM

ZHANG Jian and ZHANG Ye
College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China

Abstract At present, the employment prediction system of colleges and universities mostly adopts single traditional feature modeling, which leads to problems such as poor employment prediction effect and weak employment accurate service. This paper proposes a multi-feature fusion based on GRU-LSTM employment prediction method. Firstly, students' behavior features are added to the traditional prediction model, and the feature vector of multi-information fusion is constructed. Then, considering the different contribution of different influencing factors to college students employment, an optimal feature extraction method of employment prediction based on Pearson correlation coefficient is proposed to optimize the feature subset. Finally, a combined prediction model of GRU and LSTM is proposed, which combines the advantages of high prediction accuracy of LSTM and short prediction time of GRU to make efficient and accurate prediction of employment data. Experimental results show that compared with the traditional methods, the accuracy of employment prediction by this method increases by 4.2%, providing reliable data support for improving the employment of college students.

Keywords Deep learning, LSTM, Career prediction, Data mining

1 引言

就业问题关乎国计民生,一方面高校要对大学生就业指导、职业规划等课程设计和咨询服务不断优化和完善;另一方面,现代企业要在市场调节中不断明确人才聘用标准和专业技能水平[1-2]。但受到毕业生个体差异性和传统就业模式单一性的双重影响,高校需要借助信息化、数据化系统来搭建新的毕业生与企业供需关系。当前,虽然部分高校通过校园数字化的全面建设和大数据应用的平台搭建建立了完备的数据分析平台,并且积累了海量的学生行为信息,但是这些系统无法对大学生就业的变化趋势进行分析和研究,不能提供有价值的决策信息。因此如何对大学生就业数量进行建模与预测,对大学生就业数量进行准确分析,并为高校就业指导工作提供重要参考依据显得尤为重要[3-4]。

针对毕业生就业问题,国内学者从不同角度做过大量研究。在毕业生就业影响因素方面,Zhang[5]运用统计模型从性别、党员、英语四六级成绩等 12 个指标分析影响大学生就业的因素,并采用 Logit 模型对其进行验证。Pan 等[6]在 Zhang 的基础上加入地了对大学生性格以及未来规划的分析,更加深入总结了影响大学生就业的因素。在就业预测方面,学者们往往采用基于机器学习的方法进行毕业去向预测。Wang[7]利用支持向量机分类方法对学生消费群体进行分类,重点分析了样本的选择和函数参数选择对分类结果的影响。鉴于单一灰色模型只能描述大学生就业数量单方面变化的特点,Li[8]提出基于灰色模型和 BP 神经网络的就业预测模型,在一定程度上提高了模型预测的准确性。Zhang 等[9]在 Li[8]的基础上采用 BP 神经网络和支持向量机分别对高校毕业生就业率进行建模与预测,进而得到预测结果。综上分析,以上几类就业预测方法的研究均停留在预测模型的选择上,并没有对模型的输入特征进行分析。如就业影响因素多导致无法

基金项目:国家自然科学基金青年科学基金(61705178)
This work was supported by the Young Scientists Fund of the National Natural Science Foundation of China(61705178).
通信作者:张剑(xust-zj@xust.edu.cn)

确定影响因素对高校就业的贡献率等。针对以上问题，本文提出了一种多特征融合的大学生就业预测方法。该方法将行为特征应用到就业预测中，设计了一种新的大学生就业预测方法，对提高就业预测准确性具有一定的应用价值和重要的现实意义。

2 就业影响因素分析

影响大学生就业的因素是多层次、多维度的，应该全面、客观、真实地把握和分析。一般认为，大学生就业的个体因素往往通过学生的学习成绩、性别、政治面貌、计算机等级、外语水平、学生身份和城乡情况等特征因素来表征，而这些特征并不能准确地预测学生的就业动向，甚至会出现相同指标数据的巨大差异[10-11]。而实际调查研究表明，不同类型毕业去向的学生群体在校内的行为存在较大差异，同类型毕业去向的学生群体在校内的行为也或多或少存在差异，说明学生的行为特征在一种程度上更加真实地影响了就业的选择[12]。

因此，本文在传统特征的基础上加入学生的在校行为数据，通过对这些数据的融合分析，可以很大程度上提高预测学生毕业去向的准确性。综上，本文将影响学生就业的因素进行大类区分，分为传统因素和行为因素两部分。

(1)传统因素

传统因素主要包括学生的学习成绩、计算机等级、外语水平、学生身份等静态信息。这些数据可以通过教务处系统获取，从而直接反应学生的个人专业技能，是影响学生未来毕业去向选择的重要因素。

(2)行为因素

行为因素需要基于时间序列的有关数据呈现，比如学生每学期或每学年的各项指标数据，这是一个动态变化的过程。高校中，校园一卡通记录了在校学生各方面的数据信息，其中包括超市日常采购数据、学生食堂餐饮数据、学生宿舍门禁数据、图书馆借阅数据、教学区域学习数据、体育馆锻炼数据等。这些数据从一定程度上能间接反映出学生的行为特征。对这些行为数据进行归类和抽取可分为两类数据：一类是反应大学生经济情况的数据；另一类是反应大学生个人兴趣的数据。

1)经济情况

学生的家庭基本经济条件会影响学生毕业时的工作选择，特别是在薪酬待遇方面。然而学生的家庭经济条件涉及隐私，无法直接获取精准情况或存在获取数据虚假。但是学生在校的消费数据可以在一卡通数据中直接获取，并在一定程度上能够客观反映出学生的消费水平和经济状况。加之部分家庭困难学生能够通过在校期间的困难生认定和助学金发放进行基本数据分析，确保最终数据分析的精准性。

2)个人兴趣

图书馆借阅行为是学生在校期间的一项主要活动，学生的毕业首选去向受个人兴趣影响较大，往往择业大于就业，兴趣大于专业。大量的图书借阅数据，一方面能够反映学生对专业领域基础知识和前沿领域的学习研究情况，为基础数据做支撑；另一方面，通过各类型、各学科数据的阅读，能够用数据反映出个人阅读的习惯和偏好，同时刻画出学生个体兴趣爱好的重要特征。

3 特征向量构建

为了更加精准地对毕业生就业情况进行预测，首先将传统特征和行为特征进行量化分析，并分别采用矩阵分解和熵值判断法构造相应的特征向量。

3.1 传统特征向量构建

传统因素包括学生的课程成绩、计算机等级、英语水平等信息。本文采用矩阵分解技术，构造损失函数，求解出表征学生专业基础能力的特征向量。将计算机等级、英语水平也量化为成绩信息处理，构造一个目标矩阵。

首先对于学生的成绩信息做归一化处理，构造课程矩阵 $\boldsymbol{R}\in\boldsymbol{R}^{m\times n}$，其中矩阵 $\boldsymbol{R}$ 中的每个元素 r_{ij} 代表学生 u_i 在课程 c_j 的成绩。将矩阵 $\boldsymbol{R}$ 分解成两个矩阵 $\boldsymbol{P}_{m\times k}$ 和 $\boldsymbol{Q}_{k\times n}$ 的乘积，即有：

$$\boldsymbol{R}_{m\times n}\approx\boldsymbol{P}_{m\times k}\times\boldsymbol{Q}_{k\times n}=\hat{\boldsymbol{R}}_{m\times n} \tag{1}$$

其中，k 是一个参数，用于代表 k 种潜在的专业技能，矩阵 $\boldsymbol{P}_{m\times k}$ 代表每个学生所对应的专业技能的能力，矩阵 $\boldsymbol{Q}_{k\times n}$ 用来表示 k 种专业技能与每门课程的对应关系。

为确保真实成绩 $\boldsymbol{R}_{m\times n}$ 和重构成绩 $\hat{\boldsymbol{R}}_{m\times n}$ 之间尽可能相似，本文采用两矩阵误差的平方作为损失函数 L。

$$L=\min\sum_{i,j}I_{i,j}(r_{i,j}-\boldsymbol{p}_i'\boldsymbol{q}_i)^2+\lambda(\sum_i\|\boldsymbol{p}_i\|+\sum_j\|\boldsymbol{q}_j\|) \tag{2}$$

其中，$I_{i,j}$ 代表学生 u_i 是否修过课程 c_j；如果修过，则为 1；反之，则为 0。$\boldsymbol{p}_i$ 为矩阵 $\boldsymbol{P}_{m\times k}$ 中第 i 行的列向量，$\boldsymbol{q}_j$ 为矩阵 $\boldsymbol{Q}_{k\times n}$ 中第 j 列的向量。

利用随机梯度下降法求式(2)中的 $\boldsymbol{P}_{m\times k}$ 和 $\boldsymbol{Q}_{k\times n}$ 中元素，梯度更新公式如下：

$$\varepsilon_{i,j}=r_{i,j}-\boldsymbol{p}_i'\boldsymbol{q}_j \tag{3}$$

$$\boldsymbol{p}_i=\boldsymbol{p}_i+\alpha(I_{i,j}\varepsilon_{i,j}\boldsymbol{q}_j-\lambda\boldsymbol{p}_i) \tag{4}$$

$$\boldsymbol{q}_j=\boldsymbol{q}_j+\alpha(I_{i,j}\varepsilon_{i,j}\boldsymbol{p}_i-\lambda\boldsymbol{q}_j) \tag{5}$$

其中，α 为每步梯度更新时的幅度，λ 为正则化系数，$0<\alpha\leqslant1$。

利用上述梯度更新公式求解损失函数，得出表征学生专业技能的向量 $\boldsymbol{p}_{u,k}$ 作为特征输入代入预测模型。

3.2 行为特征向量构建

本文选取学生校园消费和图书馆借阅两类典型行为来反映学生在经济和兴趣两方面的特征。

在信息论与概率统计中，熵(Entropy)是表示随机变量不确定性的度量。学生行为规律可被视为一种行为的重复性，根据学生行为在给定时间间隔内重复发生的概率分布来计算熵，从而来量化学生行为的规律性。行为熵构造方法如下所示。

假定将周期 T 分割成 n 个时间段。

$$T=\{t_1,t_2,t_3,\cdots,t_n\} \tag{6}$$

这里根据实际情况设置不同的时间周期 T 和时间间隔 n。那么每位学生在给定时间间隔 $t_i\in T$ 内发生行为 $v\in V$ 的概率为：

$$P_v(T=t_i)=\frac{n_v(t_i)}{\sum_{i=1}^{n}n_v(t_i)} \tag{7}$$

其中，$n_v(t_i)$ 指在给定时间段 t_i 时间间隔内行为 v 发生的频次，则行为 v 在这个时间段内的熵表示为：

$$E_v=-\sum_{i=1}^{n}P_v(T=t_i)\log P_v(T=t_i) \tag{8}$$

熵值越大，学生行为 v 的概率分布越平均，则表明学生的行为越不规律。对于不规律的行为，可以减小时间间隔 n 的划分。如当计算学生的消费行为可以将时间周期设置为一周，以每天作为时间间隔进行划分，因此，两种行为向量的

表示如下：

(1)为每一位学生创建相应行为指标的时间序列，分别为：描述学生消费行为的特征向量、表征学生个人兴趣的特征向量。

$$\boldsymbol{V}=\{\text{消费行为},\text{图书馆行为}\} \tag{9}$$

(2)分别计算每种行为的熵值 E_v，判断学生的行为是否规律，进而设置不同的时间间隔对每种行为进行计数，每个时间序列都汇总不同的行为(如餐厅就餐、超市消费和图书馆借阅等)，据此得出详细的行为信息。

(3)最后对获取到的行为信息使用基于 LSTM 算法的记忆模型提取出时序行为特征。

4　基于皮尔逊分析多特征融合

为了更加精准地从众多影响就业的因素中选出对就业贡献率高的影响因素，分别将传统特征、行为特征和就业数据采用皮尔逊相关分析，根据皮尔逊相关分析得出最优子集，并将其作为预测模型的输入。

皮尔逊相关系数(PCCs)用于度量两个变量之间的线性关系，用 r 表示，系数越大，相关性越强，反之越弱。已知对于两个 n 维向量 $\boldsymbol{X},\boldsymbol{Y}$，其中 $\boldsymbol{X}=[X_1,X_2,\cdots,X_n]$，$\boldsymbol{Y}=[Y_1,Y_2,\cdots,Y_n]$，$\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}}$ 分别为变量 $\boldsymbol{X},\boldsymbol{Y}$ 的均值，则两变量的相关系数表达为：

$$r=\frac{\sum_{i=1}^{n}(X_i-\overline{\boldsymbol{X}})(Y_i-\overline{\boldsymbol{Y}})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{\boldsymbol{X}})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\overline{\boldsymbol{Y}})^2}} \tag{10}$$

综上分析，计算学生的任一特征向量 $\boldsymbol{V}=(V_1,V_2,\cdots,V_n)$ 与就业数据 $E=(E_1,E_2,\cdots,E_n)$ 的相关系数，则有：

$$R=\frac{\sum_{i=1}^{n}(V_i-\overline{V})(E_i-\overline{E})}{\sqrt{\sum_{i=1}^{n}(V_i-\overline{V})^2}\sqrt{\sum_{i=1}^{n}(E_i-\overline{E})^2}} \tag{11}$$

其中，$\overline{V},\overline{E}$ 分别为变量 $\boldsymbol{V},E$ 的均值，进而得到每一特征向量与就业数据的相关性系数矩阵为：

$$\boldsymbol{r}_z=\begin{bmatrix} \boldsymbol{r}_{z-1-1} & \boldsymbol{r}_{z-1-2} & \boldsymbol{r}_{z-1-3} & \cdots \\ \boldsymbol{r}_{z-2-1} & \boldsymbol{r}_{z-2-2} & \boldsymbol{r}_{z-3-3} & \cdots \\ \boldsymbol{r}_{z-3-1} & \boldsymbol{r}_{z-3-2} & \boldsymbol{r}_{z-3-3} & \boldsymbol{r}_{z-3-n} \\ \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{r}_{z-n-1} & \boldsymbol{r}_{z-n-2} & \boldsymbol{r}_{z-n-3}\cdots & \boldsymbol{r}_{z-n-n} \end{bmatrix} \tag{12}$$

其中，相关性系数矩阵下标 z 表示不同的类型的特征向量，即专业技能的特征向量、消费行为的特征向量、个人兴趣的特征向量、家庭经济的特征向量。各相关系数表示就业数据与特征向量的相关性，将已量化好的各项学生行为特征进行归一化处理，与学生就业去向进行相关性分析，选择相关性高的特征做为数据的输入。

5　GRU-LSTM 就业预测模型构建

毕业生就业数据系统是一个具有序列相关、非平稳、非线性特征的复杂系统。从大量毕业生的就业情况以及就业影响因素分析来来看，学生的行为特征直接影响学生的就业选择。由于学生行为规律具有重复性和动态性两个特点，传统的机器学习方法在动态重复性的行为下无法获取学生有效的行为特征和数据指标。而 LSTM 神经网络不仅能够有效预测毕业生就业数据的长期和短期动态趋势，且具有较高的预测精度。但是，LSTM 参数过多，导致训练过程要花费更多的时间。因此，本文提出一种基于 GRU-LSTM 组合的就业预测模型。

5.1　LSTM 预测模型

LSTM(Long Short-term Memory)[13-14] 由输入门、遗忘门和输出门组成。LSTM 的关键是神经元状态，LSTM 通过门结构来去除或者增加信息，包含 sigmoid 激活函数的神经网络层和乘法操作，其网络结构如图 1 所示。

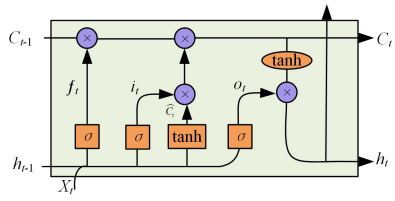

图 1　LSTM 网络结构图

Fig. 1　LSTM network structure diagram

遗忘门 f_t 决定从神经元中丢弃什么信息，该门会读取 h_{t-1} 和 x_t，输出一个在 0-1 之间的数值。如果输出值为 1，表示完全保留，0 表示完全丢弃，遗忘门的公式如下：

$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f) \tag{13}$$

输入门 i_t 决定了 x_t 中哪些新的输入可以存储在神经元中，输入门的计算公式如下：

$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i) \tag{14}$$

$$\hat{C}_t=\tanh(W_c\cdot[h_{t-1},x_t]+b_c) \tag{15}$$

$$C_t=f_t*C_{t-1}+i_t*\hat{C}_t \tag{16}$$

输出门用于控制神经元状态的输出并将状态转移到下一个神经元，输出门的计算公式如下：

$$O_t=\sigma(W_o\cdot[h_{t-1},x_t]+\boldsymbol{b}_o) \tag{17}$$

其中，$\boldsymbol{W}_o,\boldsymbol{b}_o(*=f,i,o)$ 分别为对应的权重矩阵和偏差向量。

最后，通过运行 sigmoid 函数来确定神经元状态的输出部分，神经元状态通过 tanh 进行处理后和 sigmoid 的输出相乘，最终输出部分计算公式如下：

$$h_t=O_t*\tanh(C_t) \tag{18}$$

5.2　GRU 预测模型

门控循环单元(GRU)是循环神经网络的一种，与 LSTM 结构相似[15]。GRU 将遗忘门和输入门合并为一个“更新门”。由于 GRU 减少了一个门，矩阵乘法变小，因此当训练数据量很大时，GRU 可以节省大量的时间。GRU 网络结构图如图 2 所示。

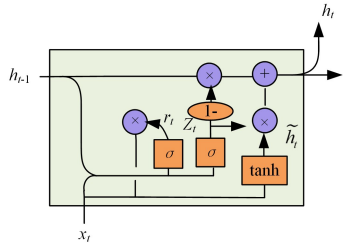

图 2　GRU 网络结构图

Fig. 2　GRU network structure diagram

图中的 r_t 和 z_t 分别表示更新门和重置门，更新门用于控制前一时刻的状态信息被带入到当前状态中的程度，更新门的值越大，说明前一时刻的状态信息带入越多。如果当前

输入之前的信息与当前的输入有一定关系则更新。更新值属于区间[0,1],由激活函数 sigmoid 决定。

$$r_t=\sigma([W_r\cdot[h_{t-1},x_t]) \tag{19}$$

重置门 z_t 的主要功能是确定有多少历史信息不能传递到下一个状态。重置门越小,前一状态的信息被写入就越少,如果当前输入之前的信息与当前的输入相关性不大或者无关,则重置。

$$z_t=\sigma([W_z\cdot[h_{t-1},x_t]) \tag{20}$$

计算出更新门和重置门后,GRU 将会计算$\tilde{h}_t$,候选隐藏状态计算公式如下:

$$\tilde{h}_t=\tanh([W_{\tilde{h}}\cdot[r_t*h_{t-1},x_t]) \tag{21}$$

最后 t 时刻 GRU 的输出 h_t 的计算公式如下:

$$h_t=(1-z_t)*h_{t-1}+z_t*\tilde{h}_t \tag{22}$$

5.3 GRU-LSTM 组合预测模型

通过综合考虑预测精度和预测时间两个方面的因素,本文提出了一种基于门控循环单元(GRU)与长短期记忆网络(LSTM)的组合预测模型对就业数据进行预测,其网络结构图如图 3 所示。

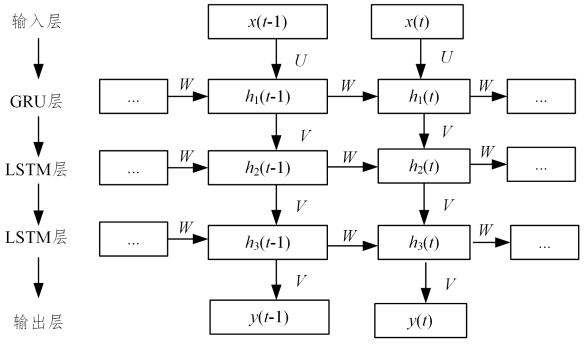

图 3 组合预测模型网络结构图

Fig. 3 Combined prediction model network structure diagram

组合模型的网络结构为三层,第一层采用 GRU,由于 GRU 的网络结构简单,参数少且易收敛,因而训练速度快,减少了训练时间,但在精度上不如 LSTM。模型的第二层和第三层均采用 LSTM,且双层 LSTM 较单层 LSTM 可获得更高的精度。因此,本文利用 GRU 训练速度快和 LSTM 预测性能好的优点,提出了 GRU-LSTM 组合预测模型,不仅确保了预测精度,也减少了预测时间,可以高效地对就业数据进行预测。

6 大学生就业预测与结果分析

本文融合了学生的传统特征和行为时序特征,使用 GRU-LSTM 组合模型构建而成。这里获取了西安科技大学 2020—2021 年 8000 名学生的就业信息,将其中 20% 的数据作为训练集,将 80% 作为测试集,得出本模型的预测结果。在评价指标的选择上选取常用的精确率(P)与召回率(R)、$F1$ 值进行评价。

$$precision=\frac{TP}{TP+FP} \tag{23}$$

$$recall=\frac{TP}{TP+FN} \tag{24}$$

$$F1=\frac{2*precision*recall}{precision+recall} \tag{25}$$

其中,TP 为结果和实际情况一致的正类数据的个数,FP 为将负样本预测为正类数据的个数,FN 为结果和实际情况不符的正类数据的数量,$F1$ 值是准确率和召回率的调和平均值,取值范围 1 到 0 代表从优到差。

6.1 特征选择

特征选择可以从原始数据特征集中选出若干个具有代表性的特征子集。因此,在模型训练之前,本文对学生的学习成绩、性别、计算机等级、外语水平、学生身份、经济情况和兴趣数据等进行皮尔逊相关分析,得到的结果如图 4 所示。

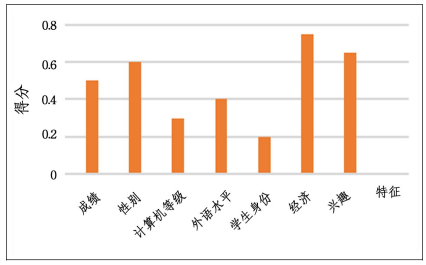

图 4 不同特征的得分图

Fig. 4 Score charts for different features

从图 4 可以看出,经济水平和兴趣是影响学生就业得分最高的两个特征,而这两个因素恰好是通过行为因素来直接反应的。因此,本文实验根据得分比例设置权重如下:

$$W=W1*0.15+W2*0.17+W3*0.09+W4*0.12+W5*0.06+W6*0.22+W7*0.19 \tag{26}$$

其中,W 为特征输入,W_1-W_7 分别对应学习成绩、性别、计算机等级、外语水平、学生身份、经济情况和兴趣数据。

6.2 模型的调参和预测

(1)训练过程分析

采用本文提出模型对就业数据进行训练,分别得到就业数据集训练及测试过程的损失函数(loss)、分类准确率(acc)随迭代次数的变化曲线. 如图 5 和图 6 所示。

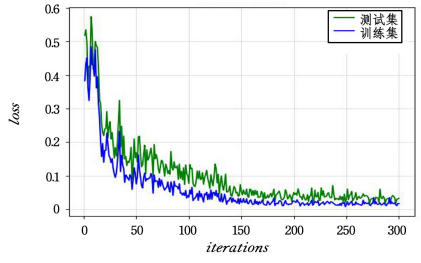

图 5 loss 函数随迭代次数的变化曲线

Fig. 5 Change curve of loss function with the number of iterations

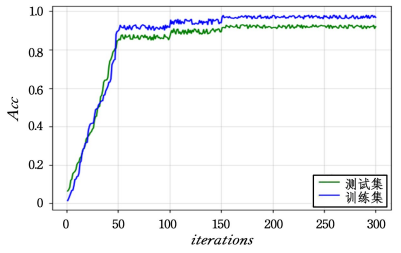

图 6 分类准确率(acc)随迭代次数变化曲线

Fig. 6 Change curve of classification accuracy with the number of iterations

可以看出,对于 $loss$ 函数,学生就业数据的训练集和测试

集呈现平稳下降趋势，均在 50 次迭代处逐渐收敛。从 150 次迭代处开始，*loss* 函数在较小值范围内变化且持续较长时间，由此可证明训练逐渐趋于稳定。对于精度值，训练集和测试集在 50 次迭代之前呈震荡上升趋势，直到 50 次迭代处精度到达较高值，在 150 次迭代后精度趋于稳定，稳定值约为 0.97，正好与 *loss* 函数在 150 次迭代处逐渐趋于稳定对应。该趋势持续一段时间后，证明训练结果已达到最优，可停止。

(2)隐藏层不同神经元个数参数分析

对第一层 GRU 隐藏层选择的神经元个数依次为 16，24，32，得出预测结果的准确率分别为 91.58%，92.96%，92.31%，训练集的准确率分别为 92.81%，95.09%，94.22%，测试集准确率分别为 93.58%，93.69%，90.32%。依次对两层 LSTM 隐藏层做相同设置，分别得出其训练集和测试集的准确率。最终确定的隐藏层神经元个数分别为 24，32，32。预测模型参数设置如表 1 所列。

表 1　预测模型参数

Table 1　Parameters of different prediction models

训练参数	参数值
第一层 GRU 隐藏神经元数量	24
第二层 LSTM 隐藏神经元数量	32
第三层 LSTM 隐藏神经元数量	32
激活函数	Tanh(x)
迭代次数	150

6.3　对比实验结果

(1)融合特征与传统特征就业预测对比实验

为了分析传统特征和融合特征对就业预测结果的影响，将传统特征和本文所提出的加入行为因素的融合特征作为模型输入，均采用 LSTM 预测模型，通过对模型进行训练和测试，最终得到预测结果表 2 和图 7 所示。

表 2　传统特征和本文特征验证结果

Table 2　Verification results of traditional features and features in this paper

输入特征	*acc*	*precision*	*recall*	*F*1
传统特征	0.625	0.648	0.659	0.630
本文特征	0.698	0.705	0.715	0.709

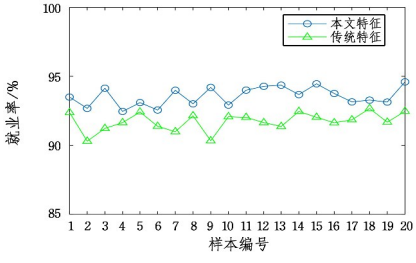

图 7　传统特征和本文特征验证结果

Fig. 7　Verification results of traditional features and features in this paper

由上述实验结果可以看出，采用多信息融合的特征向量得到的预测结果明显优于传统特征得到的预测结果。这是因为本文算法在特征的选择上加入了行为因素分析，在一定程度上更加真实地反映出学生的就业意向，因此预测结果更准确。

(2)不同模型就业预测对比实验

基于上述分析，采用多信息的融合特征能得到更好的就业预测结果。为进一步验证本文提出的就业预测模型的性能，对 GRU-LSTM 模型和单一 LSTM 模型进行十折交叉验证对比，得到的结果如表 3 和图 8 所示。

表 3　LSTM 模型和 GRU-LSTM 模型的验证结果

Table 3　Validation results of LSTM model and GRU-LSTM model

预测模型	*acc*	*precision*	*recall*	*F*1	*Time*/s
LSTM 模型	0.698	0.705	0.715	0.709	82
本文模型	0.722	0.736	0.724	0.730	58

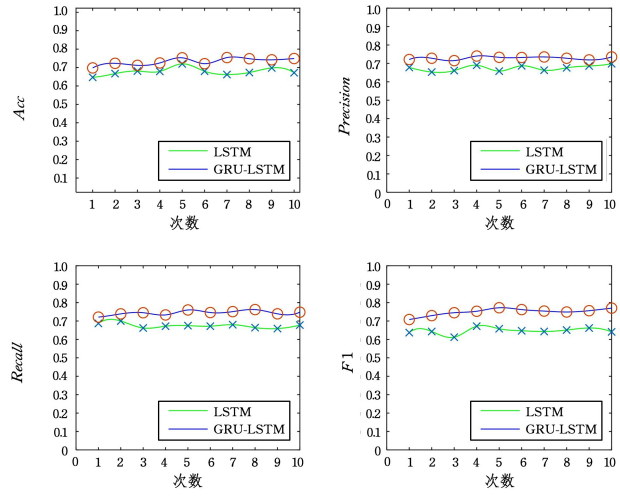

图 8　LSTM 模型和 GRU-LSTM 模型的验证结果

Fig. 8　Validation results of LSTM model and GRU-LSTM model

从上述结果可以看出本文预测模型在各个指标上都优于 LSTM 模型。因为本文模型采用非线性建模，能更好地学习并训练大量的融合特征向量，从而准确描述毕业生就业率的变化规律。同时，由于 GRU 神经网络模型作为 LSTM 的轻量级变体，在保留其拟合时间依赖性、非线性序列记忆能力的情况下减少了训练参数、降低了算法的复杂度、加快了收敛速度、提高了预测精度。经过实验研究表明，随着时间的增加，本文预测模型的 *F*1 值也会增大，预测精度也会越来越高。综上可以得出，本文提出的模型对学生的就业预测是真实有效的。

结束语　本文首先提出了一种基于传统因素和行为因素的多特征融合方法，综合考虑各个因素对就业的影响，结合皮尔逊分析选出最优特征子集，提高了数据集的质量。其次，提出了一种预测精度高和预测时间短的 GRU-LSTM 组合模型，实现了对就业数据进行高效精准预测。综上，对于高校来说，本文模型不仅可以对学生的就业情况进行预测分析，有效搭建毕业生与企业的有效桥梁，同时本模型的行为分析可以帮助培养学生的综合能力，制定科学合理的人才培养计划。对于学生来说，可为其在择业和就业的迷茫中提供优势分析，帮助其制定合理的职业规划，服务于国家现代化建设。

参 考 文 献

[1]　SUN F. Analysis of current employment situation of college students and research on measures to improve employment quality[J]. Employment and Security, 2021(13): 54-55.

[2]　KE G, MENG Q, FINLEY T W, et al. Light GBM: a highly efficient gradient boosting decision tree[C]// Neural Information Processing Systems, 2017: 3149-3157.

[3]　CHEN J T. Research and Analysis of Employment Prediction

Algorithms for College Graduates[J]. Modern Information Technology, 2019, 3(12): 86-87.

[4] ZHU Z X. Model for Relationship Between College Student Employment Category and Academic Performance[J]. Industrial Control Computer, 2021, 34(11): 108-110.

[5] ZHANG Q F. On the Comprehensive Factors that Affect College Students' Employment: An Empirical Analysis Based on the Logit Model[J]. Journal of Hunan University of Science & Technology (Social Science Edition), 2014, 17(3): 175-180.

[6] PAN Z S, YAN C. Survey of college students' employment situation under the guidance of employment quality improvement[J]. Heilongjiang Researches on Higher Education, 2019, 37(12): 139-42.

[7] WANG D C. Research and Application of Data Mining in Campus card Consumption Behavior Analysis[D]. Harbin: Harbin Engineering University, 2010, 23-36.

[8] LI X. Research on modeling and forecasting of college students employment[J]. Modern Electronics Technique, 2017, 40(21): 110-111.

[9] ZHANG Z H, LIU Z Q. Research on college graduate employment rate prediction based on big data integration technology [J]. Modern Electronics Technique, 2021, 44(4): 80-82.

[10] ZHANG X. Prediction of the Career Development Direction of College Students Based on Convolutional Neural Networks[D]. Changchun: Northeast normal university, 2020. 11-22.

[11] ZHANG K S, WANG X Q. The Development Trend of Economic New Normal: From Quantity Concept to Quality Management—Based on the Perspective of Employment Quality Analysis of College Graduates[J]. Guangdong Social Sciences, 2016(1): 36-45.

[12] YU W H. Employmentand Entrepreneurship Guidance System for College Students Based on Big Data[J]. Microcomputer Applications, 2021, 37(9): 37-39.

[13] CHEN W, CHEN J X, JIANG Y Q, et al. Fault idedntification of rolling bearing based on RS-LST[J]. China Sciencepaper, 2018, 13(10): 1134-1141.

[14] KUMAR J D, ZHANG Z, KAIQI H, et al. Multi angle optimal pattern-based deep learning for automatic facial expres sion recognition[J]. Pattern Recognition Letters, 2017: 1-9.

[15] YANG C X, HAN W, GAO Z Q. Short-term forecasting for solar irradiance using GRU neural network[J]. China Sciencepape", 2020, 15(1): 8-14.

ZHANG Jian, born in 1988, Ph.D candidate, associate professor. His main research interests include big data maintenance and analysis.