

文章编号: 2096-1618(2018)04-0365-05

脑电波情绪分类算法的研究

马新斐, 刘志宏, 姜添浩

(成都信息工程大学电子工程学院, 四川 成都 610225)

摘要:人们的情绪变化本质是大脑皮层上的高级神经活动。情绪认知应用是未来重要的一个趋势,现在以脑机接口为主流工具的研究,在脑电情绪主观世界和信号客观世界之间建立了桥梁。使用多种分类器来对情绪识别,选择有监督机器学习的 Fisher、贝叶斯、SVM 和无监督机器学习的 DBN 分类器进行研究。结果表明:在分类精度上,贝叶斯要优于 Fisher, DBN 要优于 SVM,在运行时间上,贝叶斯运行时间最短。DBN 有更高的分类精度和更低的标准偏差,平均最佳分类精度为84.01%,最低标准偏差为9.74%,比较适合脑电情绪识别。

关键词:脑电;情绪认知;分类精度;机器学习

中图分类号:TP391.4

文献标志码:A

doi:10.16836/j.cnki.jcuit.2018.04.003

0 引言

脑电波研究是探索人体生理信号的重要途径,而且脑电波的情绪认知研究也是未来智能化的一个新方向。随着医学物联网智能信息处理的发展,脑电情绪分析和应用开始成为热点问题,它的潜在前景也越来越多地进入人们的视野。人们的情绪变化本质是大脑皮层上的高级神经活动。情绪认知应用是未来重要的一个趋势,现在以脑机接口为主流工具的研究,在脑电情绪主观世界和信号客观世界之间建立了桥梁^[1-2]。

使用多种分类器来对情绪识别,选择有监督机器学习的 Fisher、贝叶斯、SVM 和无监督机器学习的深度信念网络(DBN)分类器,通过提取好的脑电特征与未经训练的特征样本经过学习和分类来确定各种情绪状态,从而达到情绪识别的目的。

1 识别算法研究

1.1 朴素贝叶斯算法

定义在贝叶斯回归中目标 t , 特征向量 x 与高斯白噪声 n 线性相关^[3], 即:

$$t = \omega^T x + n$$

在给定的假设之后,给出在回归中使用权重的似然函数:

$$p(D | \beta, \omega) = \left(\frac{\beta}{2\pi} \right)^{\frac{N}{2}} \exp \left(-\frac{\beta}{2} \| X^T \omega - t \|^2 \right)$$

其中 t 表示目标回归向量, X 为训练的特征向量水平叠加形成的矩阵, D 为 $\{X, t\}$, β 为噪声的逆方差, N 为训练样本的数量。在分类器算法中执行已经指定的变量的先验分布,表达式为:

$$p(\omega | \alpha) = \left(\frac{\alpha}{2\pi} \right)^{\frac{N}{2}} \left(\frac{e}{2\pi} \right)^{\frac{1}{2}} \exp \left(-\frac{1}{2} \omega^T I'(\alpha) \omega \right)$$

由于先验和似然都符合高斯分布的,后验也是符合高斯的,参数可从构建矩阵,从似然和先验中分离出来。 m 和协方差 C 满足:

$$m = \beta(\beta X X^T + I'(\alpha))^{-1} X t$$

$$C = (\beta X X^T + I'(\alpha))^{-1}$$

通过与似然函数相乘形成一个新的输入向量,然后和后验分布乘以 ω 的积分,得到预测分布,则输入向量为条件的回归目标的概率分布如下:

$$p(\hat{y} | \beta, \alpha, \chi, D) =$$

$$\int p(\hat{t} | \beta, \alpha, \chi, \omega) p(\omega | \beta, \alpha, D) d\omega$$

在设定中,类的概率可以通过计算训练过程中的目标概率来获得,通过上面预测分布的公式和带入 β , σ , D , α 的条件,可以得到^[4]:

$$p(\hat{y}=1 | \chi) = \frac{p\left(\hat{t} = \frac{N_1}{N} | \chi\right)}{p\left(\hat{t} = \frac{N_1}{N} | \chi\right) + p\left(\hat{t} = \frac{-N_2}{N} | \chi\right)}$$

后验分布和预测分布取决于参数 α, β 。由上述参数已知,实际情况是未知的,可以用交叉验证来预测参数产生的性能,也可以用贝叶斯回归框架提供的用似然函数,然后最大化可能性,这样更省时。

1.2 深度信念网络

深度信念网络(DBN)是深层架构的概率生成模

型,它的特点是输入数据分布使用隐藏变量,由若干层受限的 Boltzmann 机(RBM)和一层分类器组成,这里采用 BP 分类器。受限的 Boltzmann 机只有两层神经元,一层叫显层,由显元组成,一层叫隐层,由隐元组成。显层和隐层双向连接,DBN 是通过将预定先定义数量的 RBM 相互堆叠彼此之上来构建的^[5]。

DBN 在训练提取数据的流程中可以分为两部分,第一部分是由多层 RBM^[6]构建,将特征输入第一层 RBM,用无监督式学习对其进行训练,训练结果作为第二层 RBM 的输入层,每次训练一层,依次迭代。第二部分是由 BP 网络构建,它处于最末一层,用有监督式学习微调架构每一层的分类器参数,使用最后一层的隐藏层输出与给定的输入进行分类。

在一个 RBM 中,假设显元 v 和隐元 h ,模型参数为 θ ,联合分布 $p(v, h; \theta)$, 能量函数 $E(v, h; \theta)$ 则可得^[7]:

$$P(v, h; \theta) = \frac{\exp(-E(v, h; \theta))}{Z}$$

$$Z = \sum_v \sum_h \exp(-E^Z(v, h; \theta))$$

$$P(v; \theta) = \frac{\sum_h \exp(-E(v, h; \theta))}{Z}$$

Z 为正常因子, $P(v; \theta)$ 为边缘概率,显层符合高斯分布,隐层符合伯努力分布,能量函数被定义为

$$E(v, h; \theta) = -\sum_{i=1}^I \sum_{j=1}^J \omega_{ij} v_i h_j - \frac{1}{2} \sum_{i=1}^I (V_i - b_i)^2 - \sum_{j=1}^J a_j h_j$$

其中 ω_{ij} 是显元 v_i 和隐元 h_j 之间的相互作用, b_i 和 a_j 是偏差项, I, J 是显元和隐元的个数,条件概率可以被计算如下:

$$p(h_j = 1 | v; \theta) = \sigma\left(\sum_{i=1}^I \omega_{ij} v_i + a_j\right)$$

$$p(v_i = 1 | v; \theta) = N\left(\sum_{j=1}^J \omega_{ij} h_j + b_i, 1\right)$$

已知 $\sigma(x) = 1/(1 + \exp(-x))$, 显元服从高斯分布,当所得值大于一个随机值时方差取 1,输入数据与重建数据相差,越来越小,更新权值 ω 则:

$$\Delta \omega_{ij} = E_{\text{data}}(v_i h_j) - E_{\text{model}}(v_i h_j)$$

$E_{\text{data}}(v_i h_j)$ 为在训练中要分类的集合即样本期望集合, $E_{\text{model}}(v_i h_j)$ 为已经定义好的集合即可视期望集合。 $E_{\text{model}}(v_i h_j)$ 很难计算,所以选用随机采样方法,给出一个近似值。选用对比散度来代替吉布斯采样,来监督梯度变化^[8]。每一次梯度变化,吉布斯采样需要多次重复才完成,算法运行时间长,对比散度可以提高效率。

1.3 支持向量机

支持向量机(SVM)是一种二类分类模型,也可以

向多类推广,基本模型定义为特征空间上间隔最大的线性分类器^[9],学习策略是利用最远的分类间隔使两类完成分类^[7-11]。

设定训练数据: $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \mathbf{x} \in R^n, y \in \{+1, -1\}$, \mathbf{x} 为输入向量, y 为样本特征。

把样本分在一个超平面: $(\mathbf{w}^T \mathbf{x}) + b = 0$, \mathbf{w} 为权值向量, b 为偏置(负阈值)则有:

$$(\mathbf{w}^T \mathbf{x}_i) + b > 0 \quad y_i = +1$$

$$(\mathbf{w}^T \mathbf{x}_i) + b < 0 \quad y_i = -1$$

支持向量机主要是找到一个与最近样本点之间间隔(分离边缘)最大的超平面,确定 i 最大时的 \mathbf{w} 和 b 。根据已知几何知识可以得到最优距离为^[10,12]:

$$r = \frac{\mathbf{w}_0^T \mathbf{x} + b_0}{\|\mathbf{w}_0\|}$$

已知最小化代价函数: $f(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 = \frac{1}{2} \mathbf{w}^T \mathbf{w} +$

$C \sum_{i=1}^I \varepsilon_i$, C 是选定正参数,采用 Lagrange 系数方法解决约束最优问题,函数为:

$$L(\mathbf{w}, b, a) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{i=1}^I a_i [y_i (\mathbf{w}^T \mathbf{x}_i + b) - 1 + \varepsilon_i]$$

对函数求偏导,令导数为 0,可得

$$\mathbf{w} = \sum_{i=1}^I a_i y_i \mathbf{x}_i, \sum_{i=1}^I a_i y_i = 0$$

得到非线性可分数据对偶问题即:

$$Q(a) = \sum_{i=1}^I a_i - \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^I a_i a_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j$$

引入 Kuhn Tucker 最优条件^[11]

$$a_i [(\mathbf{w}^T \mathbf{x}_i + b) y_i - 1 + \varepsilon_i] = 0$$

最终推导得到最优解 \mathbf{W} 和 b , 以及最优分类判别函数:

$$f(x) = \text{sgn} \left[\sum_{i=1}^I a_i y_i (\mathbf{x}_i^T \mathbf{x} + b_0) \right], b_0 = 1 - \mathbf{w}_0^T \mathbf{x}_s$$

2 数据结果分析

分析 4 个分类器的情绪识别性能,有人监督的 Fisher 线性分类(FLDA)和朴素贝叶斯分类器(BLDA),有监督方式的支持向量机(SVM)和无监督方式的深度信念网络(DBN)。为更好比较算法的性能,将使用 BLDA 和 FLDA 进行对比分析,SVM 与 DBN 进行对比分析。

在得到降维后的特征,需要把特征集合输入分类器中进行学习,与测试数据进行测试。本实验每次送入的分类数据为 1000 个,600 个测试数据,特征维度为 32 维,分类的情绪为积极情绪,消极情绪和中性情绪。

FLDA 与 BLDA 都是有监督方式,分类在几分钟之内可以完成,不需要用户来调整参数,它们是在每次特征放入训练,挑一组实验作为验证,这段数据进行训练 20 次处理,然后分类后,其他数据来对比这组数据分类。表 1 ~ 表 4 为志愿者的 BLDA 识别率。

表 1 四通道下节律的分辨率及平均标准差

特征	Delta(Δ)	Theta(θ)	Alpha(A)	Beta(B)	Gama(γ)	全频段
PSD	51.38/14.22	48.39/9.04	57.97/15.86	64.28/14.31	66.60/15.90	74.09/15.73
DE	47.84/11.47	48.52/9.19	57.62/16.70	69.90/15.29	69.10/14.87	82.87/10.93.
DASM	42.24/5.79	40.56/7.20	45.75/9.20	46.36/12.45	47.50/12.48	70.00/14.85
RASM	41.75/5.90	40.14/6.99	45.54/9.03	45.94/12.12	47.68/12.63	69.33/14.75

表 2 六通道下节律的分辨率及平均标准差

特征	Delta(Δ)	Theta(θ)	Alpha(A)	Beta(B)	Gama(γ)	全频段
PSD	55.01/14.50	56.21/10.60	63.21/14.80	72.59/15.08	74.36/15.37	70.53/14.70
SE	37.42/7.72	41.14/7.22	43.60/9.60	45.41/10.01	45.32/11.08	71.65/14.20
DE	49.99/12.77	57.60/12.26	66.02/13.01	75.85/13.95	75.28/14.28	85.13/9.53
DASM	42.11/9.86	44.22/9.43	53.26/12.62	60.32/12.90	61.17/14.04	74.17/11.09
RASM	41.26/9.37	44.32/9.62	53.12/9.44	61.32/14.21	62.68/14.21	73.15/12.34

表 3 九通道下节律的分辨率及平均标准差

特征	Delta(Δ)	Theta(θ)	Alpha(A)	Beta(B)	Gama(γ)	全频段
PSD	59.94/15.43	59.42/11.51	70.26/13.27	78.76/11.87	79.07/13.97	69.47/14.43
SE	37.42/7.72	41.83/7.22	43.26/9.73	45.51/10.01	45.32/11.09	71.65/14.25
DE	55.32/14.78	60.26/12.60	71.60/14.20	80.20/10.21	81.22/11.27	84.04/10.35
DASM	44.11/10.94	47.22/10.53	58.26/10.29	69.32/15.62	66.17/15.83	76.17/10.23
RASM	44.26/7.72	45.32/10.37	58.12/10.90	66.32/16.23	65.68/14.64	79.15/10.82

表 4 十二通道下节律的分辨率及平均标准差

特征	Delta(Δ)	Theta(θ)	Alpha(A)	Beta(B)	Gama(γ)	全频段
PSD	57.94/14.68	62.42/13.80	65.66/15.82	75.76/12.87	75.57/12.97	62.47/15.40
SE	37.42/7.50	41.14/7.22	44.26/9.76	45.51/9.99	45.32/11.08	71.65/14.28
DE	55.32/13.55	62.26/13.70	68.60/15.01	77.85/13.60	77.22/14.35	86.64/8.63
DASM	47.11/10.42	46.22/9.92	59.26/12.38	69.32/15.19	70.17/14.43	76.17/14.06
RASM	45.26/12.21	45.97/10.29	59.12/11.15	69.32/15.30	70.68/13.67	75.15/13.74

所有的实验组里,除了第 3 组和第 7 组的分类精确度小于平均分类 80%,其他组都在之上。第 3 组被报告,在实验期间,刺激材料有一次错误的播放,除去错误,3 组也高于平均分类率。第 7 组比较低的原因可能在于志愿者实验时间太久,出现疲劳状态,第 1 组表现最好,因为受试者高度集中和积极的意愿。

比较 FLDA 与 BLDA 的两种分类方法,选用 4 个电极,8 个电极,16 个电极,32 个电极组成的电极配置,来对比它们的分类精度。超过 8 个电极的情况下,BLDA 的分类准确率开始提升,FLDA 的分类准确率开始下降,由 4 个电极的情况下,BLDA 的分类准确率稍微高于 FLDA 的准确率。超过 8 个电极,BLDA 的分类准确率要明显高于 FLDA,在 32 个全电极情况下,BLDA 和 FLDA 的分类率都比较高,最好的分类效果达到

85%。可得,在 16 电极或者 32 电极使用 BLDA 可进一步提高识别率,在不超过 8 电极下可以选用 FLDA。这是因为 FLDA 算法无法处理特征数过大的训练数据实例,当样本发生过大时,FLDA 中的全类离散度矩阵变为奇异。BLDA 既可以解决样本少的问题,更通过正则化解决样本重复问题,效率高不需要用户干预,也不用费时的交叉验证,有良好的分类精度。

DBN 分类器是无监督分类器。SVM 的基本思想是将输入特征投入由核心传递函数放入更高维度的特征空间,这样要比原特征空间更好分离。SVM 采用线性内核,搜索的空间参数定为 $2^{[-10:10]}$, $C=1$,来找最佳值。对于 DBN,可构建两个隐层,在第一层和第二层隐层搜索最佳神经元,隐层节点数分别为 $[200:500]$, $[150:500]$,设置无监督学习率和监督学习率分别为

0.5和0.6。防止在局部最小值卡住,用动量来更新权值。在将特征放入 DBN 前,特征值会被缩放 0 ~ 1 倍,除以标准差,最后加0.5。

选取从 4 组实验中选取 30 个实验数据,选取的特

征为不同特征,特征来自不同节律带,用分类器对这些数据的分类精度和标准偏差来对比分类器,具体数据见表 5。从中可以证实在 Gamma 和 Beta 频段,脑电信号更活跃,这与上述发现一致。

表 5 不同分类器分类率

实验序号	分类器	Delta(Δ)	Theta(θ)	Alpha(α)	Beta(β)	Gama(γ)	全频段
1	SVM	62.65	68.42	93.62	97.45	97.57	94.47
	DBN	65.42	65.33	95.97	98.01	98.32	96.65
2	SVM	47.47	53.65	55.47	63.02	84.38	77.08
	DBN	67.97	57.57	61.40	73.21	73.53	78.37
3	SVM	64.89	60.43	63.83	79.12	79.57	84.26
	DBN	77.90	56.53	55.16	84.03	86.13	89.03
4	SVM	54.17	77.34	91.78	96.61	91.93	90.89
	DBN	57.90	72.05	79.53	95.47	97.67	93.68
标准分类率	SVM	56.43	59.57	66.59	80.97	83.86	83.08
	DBN	61.98	59.83	67.27	81.49	84.05	84.01
标准差	SVM	7.70	11.02	18.84	14.84	14.36	10.66
	DBN	8.36	6.69	13.40	10.16	10.84	9.74

从表 5 可得,DBN 在全频段对 DE 特征分类最佳,从结果可以看到,最佳分类精度为96.65%,最低标准偏差为7.62%。而对于 SVM,结果相似,在所有节律带,DE 特征分类最佳。结果表明,与其他特征相比,DE 的性能更优越,同时因为左右脑区非对称特征(DASM,RASM)的维数要少于 PSD 和 DE 特征,分类精度也比较接近,可以证明不同脑区的活动对情感识别是有影响的。

图 1 是 4 种分类器的平均标准差。结果表明,DBN 的分类平均标准差最小,即误差最小,深度学习方法有更好的选择能力,过滤掉不相关的特征,实现更好的分类准确性。特征提取和特征选择在情绪识别中至关重要,DBN 可以有效结合有监督和无监督学习来进阶特征提取和特征选择。分析结果表明,SVM 积极情绪分类精度比较高,消极情绪很难预测,在 SVM 中,消极情绪和平静情绪容易混淆,不能很好地分类消极情绪。

3 结束语

通过两种类别的分类器,根据特征分别对脑电情绪进行分类,分为 3 种情绪,消极的、积极的、平静的,在无人监督学习中,FLDA 和 BLDA 都可以成功地对情绪分类,但是 FLDA 处理不了特征数过大的数据,对于 32 导联通道,BLDA 有更好的识别率。对于有监督学习,SVM 和 DBN 对于情绪分类效果要比无监督的 FLDA 和 BLDA 分类精度高,其中 DE 特征分类的效果最好,DBN 的分类精度均值和标准差要高于 SVM,SVM 在分类过程中,很难区分消极情绪和平静情绪的差异,DBN 比较适合脑电情绪分类。所以 DBN 比 SVM 有更高的分类精度和更低的标准偏差,更适合脑电情绪识别。

参考文献:

[1] 聂聿,王晓辉,段若男,等. 基于脑电的情绪识别研究综述[J]. 中国生物医学工程学报,2012,31(4):595-606.

[2] 刘广权. 无监督自适应式脑机接口的算法研究[D]. 上海:上海交通大学,2010.

[3] 吴婷,颜国正,杨帮华,等. 基于有监督学习的概率神经网络的脑电信号分类方法[J]. 上海交通大学学报,2008,42(5):803-806.

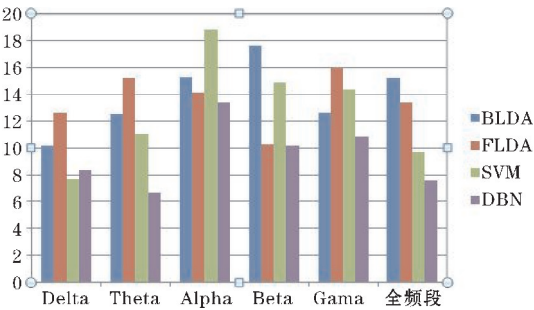


图 1 4 种分类器的平均标准差

- [4] 史原. 基于 Fisher 判别法的脑电图数据分析的研究[J]. 微计算机信息, 2010, 26(7): 226–228.
- [5] Yang G, Lin Y, Bhattacharya P. A driver fatigue recognition model based on information fusion and dynamic Bayesian network [J]. Information Sciences, 2010, 180(10): 1942–1954.
- [6] Wang Z F, Wang Z H. An Optimization Algorithm of Bayesian Network Classifiers by Derivatives of Conditional Log Likelihood[J]. Chinese Journal of Computers, 2012, 35(2): 364–374.
- [7] Kummer B, Schultz R. Kuhn-Tucker-points of parametric convex programs as solutions of perturbed generalized equations [J]. Revista Brasileira De Anestesiologia, 1986, 81: 28–36.
- [8] Cummer J. Methodology and Techniques for Building Modular Brain-Computer Interfaces [J]. Machine Learning, 2014.
- [9] Guo Y, Wang S, Gao C, et al. Wishart RBM based DBN for polarimetric synthetic radar data classification[C]. IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2015: 1841 – 1844.
- [10] 高琰, 陈白帆, 晁绪耀, 等. 基于对比散度-受限玻尔兹曼机深度学习的产评论情感分析[J]. 计算机应用, 2016, 36(4): 1045–1049.
- [11] 王凯, 侯著荣, 王聪丽. 基于交叉验证 SVM 的网络入侵检测[J]. 测试技术学报, 2010, 24(5): 419–423.
- [12] Kousarrizi M R N, Ghanbari A R A, Teshnehlal M, et al. Feature Extraction and Classification of EEG Signals Using Wavelet Transform, SVM and Artificial Neural Networks for Brain Computer Interfaces [C]. International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing. IEEE, 2009: 352–355.
- [13] Górecka J. Removing physiological artifacts from the EEG data by algorithms based on differential entropy[J]. Pomiar Automatyka Kontrola, 2012, 58: 975–977.
- [14] Zheng W L, Guo H T, Lu B L. Revealing critical channels and frequency bands for emotion recognition from EEG with deep belief network [C]. International Ieee/embs Conference on Neural Engineering. IEEE, 2015: 154–157.
- [15] 刘冲, 赵海滨, 李春胜, 等. 脑电信号频带能量特征的提取方法及分类研究[J]. 系统仿真学报, 2012, 24(12): 2496–2499.

Study of Brain Wave Emotion Classification Algorithm

MA Xin-fei, LIU Zhi-hong, JIANG Tian-hao

(College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: The emotion belongs to higher nervous activity in the cerebral cortex of human. Now many researchers use BCI in formal analysis, simulation, and phototyping to explore predicted system behavior between the subjective world of emotion and the objective world of the signal. This paper compares various classifiers of emotion recognition, and then applies two sets of classifiers. The unsupervised classification include DBN, the supervised classification include Bayes classifier and Fisher classifier and SVM. The results have shown that the DNB method performed better than SVM in classification accuracy, and the Bayes classifier is better than Fisher classifier in running time. DBN has a higher classification accuracy and lower standard deviation, and it is more suitable for EEG emotion recognition. Moreover, the average classification accuracy is 84.01% and the minimum standard deviation is about 9.74%.

Keywords: EEG; emotion recognition; classification accuracy; machine learning