

北京化工大学

Beijing University Of Chemical Technology

基于小波变换和支持向量机的 癫痫脑电信号分类

期末答辩









本研究基于**离散小波变换**、三种机器学习算法 (**支持向量机、K最近邻算法和决策树**)以及集成 学习中的stacking方法,使用matlab和python 两个工具,研究了两个数据集不同情况下的癫 痫脑电信号分类问题。



目录 CONTENTS

研究背景Research Background

研究内容Research Content

研究结果 Research Result 研究结论 Research Conclusion

研究背景

Research Background



北京化工大学



② 诊断困难

- ◆ 昏厥、偏头疼等疾病可能与癫痫有类似的表现和病征
- ◆ 临床诊断主观性较强,可能会出现误诊,同时效率较低

○ 意义重大

- ◆ 及时诊断可减少对患者的不利影响
- ◆ 及时诊断可减轻患者家庭的经济压力



研究内容

Research Content



北京化工大学





波恩大学癫痫脑电数据集

- ◆ 五组脑电数据(A~E)
- ◆ 采样频率173.61Hz
- ◆ 每组100个脑电片段,格式为txt
- ◆ 每个片段23.6s, 4097个数据点
- ◆ 已进行带通滤波(0.53~40Hz)

情况	志愿者临床状态	EEG子集
1	正常 vs 发作期	A vs E
2	正常 vs 发作期	B vs E
3	发作间期 vs 发作期	C vs E
4	发作间期 vs 发作期	D vs E

表2.1 四种分类任务



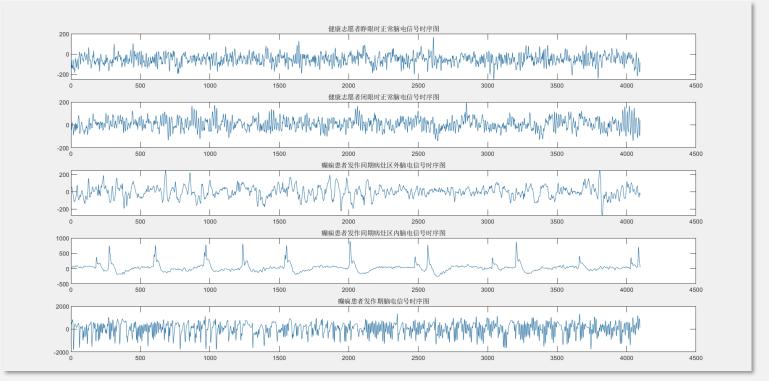


图2·1 脑电信号时序图(A-E)







CHB-MIT脑电数据集

- ◆ 病例chb01, chb03
- ◆ 采样频率256Hz
- ◆ 每个病例发作间期和发作期的时间 均挑选**240s**,格式为edf
- ◆ 选择FT9-FT10通道
- ◆ 带通滤波(0.1~60Hz)

病例	患者临床状态
chb01	发作间期 vs 发作期
chb03	发作间期 vs 发作期

表2.2 两种分类任务



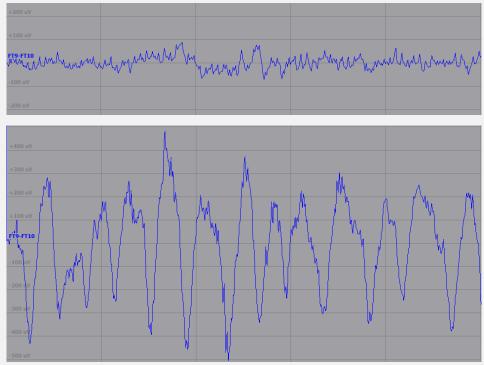
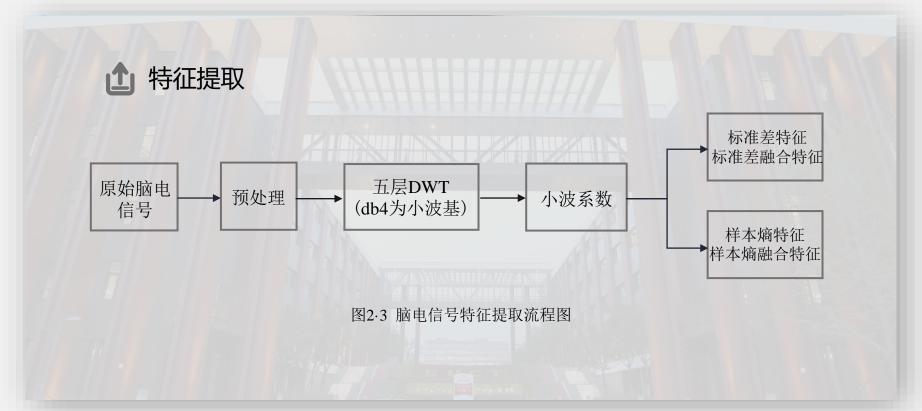


图2·2 病例chb01发作间期(上)与发作期(下)脑电信号时序图









研究内容

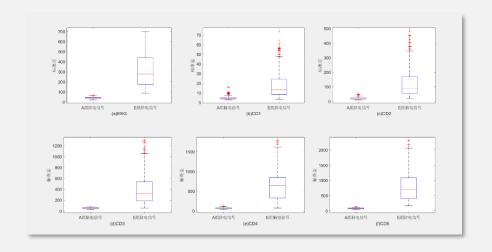


图2·4 AE两组标准差箱型图对比

图2·5 病例chb01发作间期和发作期 脑电信号方差箱型图对比





		特征数			
数据集	分类情况	标准差特征	标准差	样本熵特征	样本熵
			融合特征		融合特征
波恩大学癫痫 脑电数据集	A vs E	6	1	4	1
	B vs E	6	1	4	1
	C vs E	6	1	4	1
	D vs E	6	1	4	1
CHB-MIT脑电 数据集	病例chb01	6	1	3	1
	间期vs发作期				
	病例chb03	6	1	2	1
	间期vs发作期				

表2.3 不同分类情况特征数汇总





沙 支持向量机

- ◆ 基于单一特征的二分类,采用默认的**线** 性核函数
- ◆ 基于融合特征的二分类,采用RBF径向 基核函数
- ◆ 分别根据单一特征与融合特征对不同情况进行分类
- ◆ 进行五折交叉验证,结果取均值

情况	cost	gamma
A vs E	16	22.6274
B vs E	11.3137	32
C vs E	22.6274	32
D vs E	32	32

表2·2 样本熵融合特征二分类网格搜索最优c、g





≪ K最近邻算法

- 使用欧几里得距离进行度量
- 最近邻参数K经过比较选取为9
- 分别根据单一特征与融合特征对不同 情况进行分类
- 进行五折交叉验证, 结果取均值

决 决策树

- 使用CART算法生成二叉树
- 设置决策树的自动剪枝为'on'
- 分别根据单一特征与融合特征对不 同情况进行分类
- 进行五折交叉验证, 结果取均值



◎ 集成学习的stacking方法

- ◆ 本研究中, stacking方法由两层模型构成
- ◆ 第一层模型由**支持向量机和K最近邻算法**构成,第二层模型由**决策树**构成
- ◆ 算法参数与上述参数一致
- ◆ 只根据融合特征对上述两数据集不同情况进行分类
- ◆ 进行五折交叉验证,结果取均值





房 总结

- ◆ 本研究以'db4'为小波基,对预处理后的数据进行了**五层**离散小波变换,之后根据小波系数计算得出标准差特征与样本熵特征,并通过绘制箱型图对比, 选取得到要用的特征
- ◆ 本研究使用支持向量机、KNN和DT三种分类算法与集成学习的stacking方法,分别根据"表2·3"中的不同特征进行分类,最后选用**准确率、灵敏度**和**特异度** 三个评价指标对分类结果进行评价

研究结果

Research Result



北京化工大学



作者		本文		Zhu G etc ^[21] (2013)	Samiee K etc [22] (2015)	Siuly S etc ^[23] (2019)
分类算法	SVM	KNN	DT	Multi Scale K-Means	MLPC	KNN
A vs E	99.67%	100%	100%	100%	99.8%	100%
B vs E	99.79%	99.54%	99.42%	99.0%	99.3%	100%
C vs E	98.88%	98.29%	97.96%		98.5%	100%
D vs E	97.04%	96.63%	96.21%		94.9%	100%

表3.1 与同类研究分类准确率对比结果

研究结果

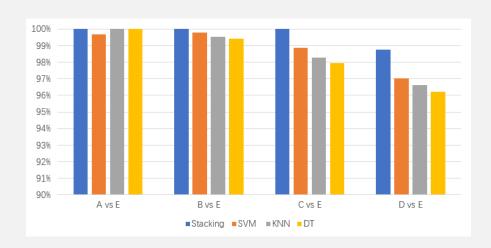


图3·1 stacking方法与三种分类算法根据标准 差特征对四种情况分类准确率比较

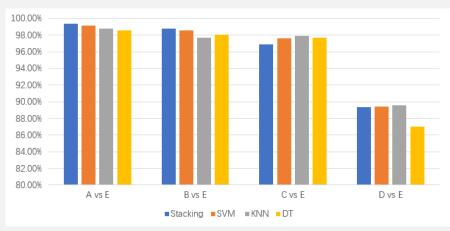


图3·2 stacking方法与三种分类算法根据标准 差特征对四种情况分类准确率比较

研究结果

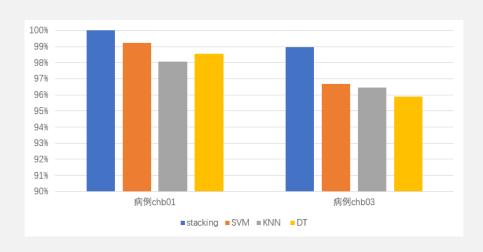


图3·3 stacking方法与三种分类算法根据标准 差特征对两个病例分类准确率比较

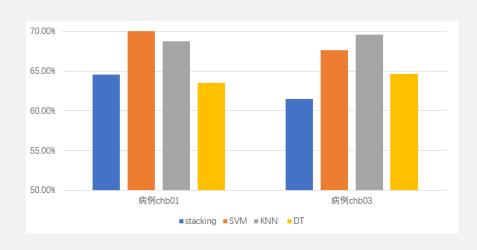


图3·4 stacking方法与三种分类算法根据标准 差特征对两个病例分类准确率比较



研究结论

Research Conclusion



北京化工大学



- ◆ 使用SVM、KNN和DT对"表2·1"中四种情况分类结果表明标准差特征优于样本熵特征,对"表2·2"的分类结果验证了这一点
- ◆ 对"表2·1"的整体分类结果表明, **SVM**分类效果最好; 对"表2·2"的分类结果表明SVM和 KNN这两种分类算法分类效果均很好
- ◆ 对"表2·1"和"表2·2"的分类结果表明,根据融合特征进行分类得到的分类准确率要高于 单一特征
- ◆ "表3·1"中与前人研究成果的对比表明,本研究提出的方法可以较好地完成癫痫脑电信号的分类任务
- ◆ stacking方法比SVM、KNN和DT这三种算法相比,大多数情况下可以提高分类准确率, 具有一定的泛化能力。





北京化工大学

Beijing University Of Chemical Technology

感谢各位老师批评指正

Thank you for the criticism of the teachers





