

基于深度学习的行为识别数据分析

—— “深度学习入门及 Python 训练” 课程报告

龚梓阳

更新：2021 年 6 月 7 日

1 数据介绍



图 1: 指定时间内完成指定行为（步行）的智能手机传感器数据记录 [1]

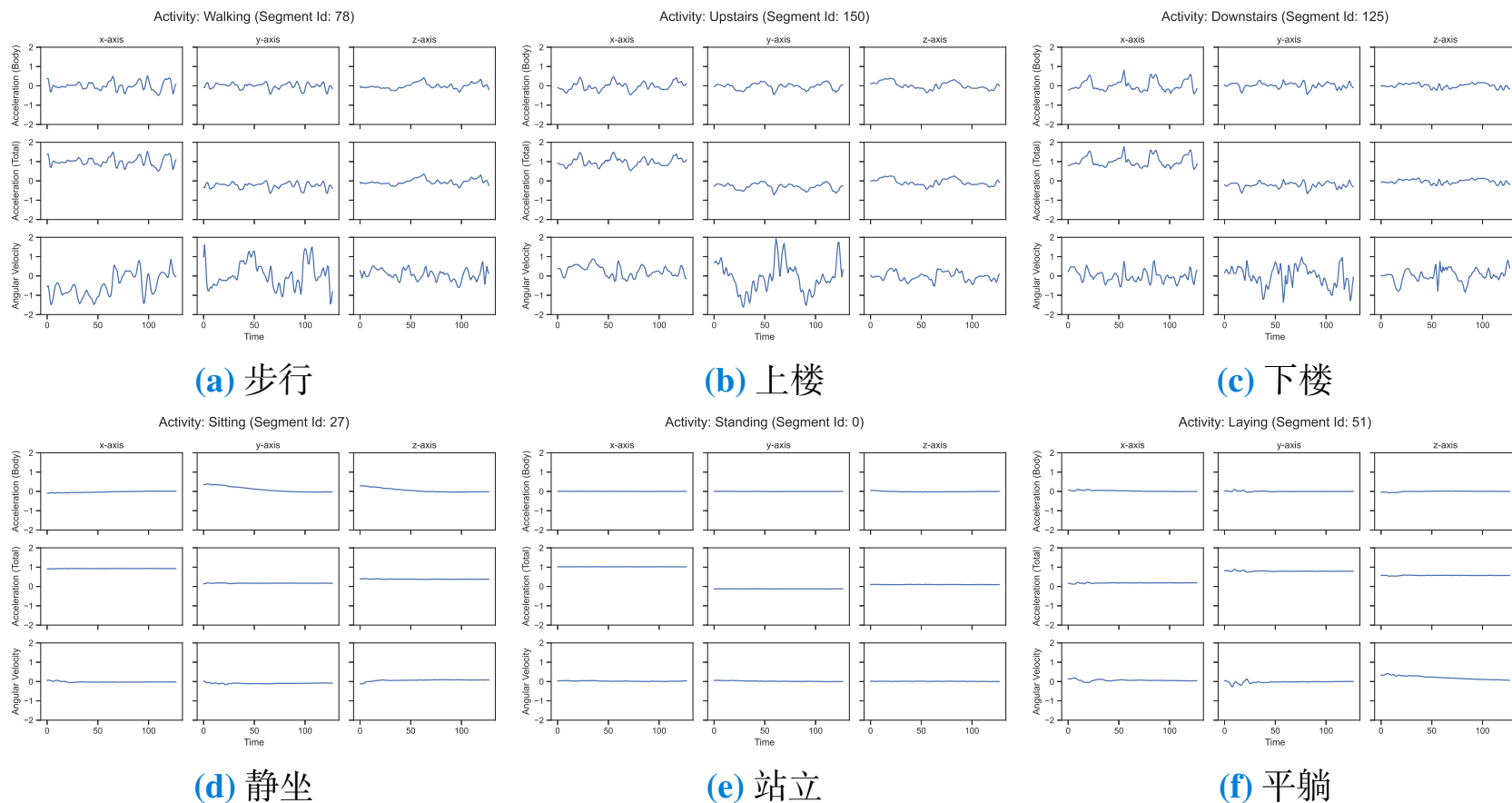


图 2: 指定时间内完成不同指定行为（步行，上楼，下楼，静坐，站立和平躺）的智能手机传感器数据记录

2 模型构建

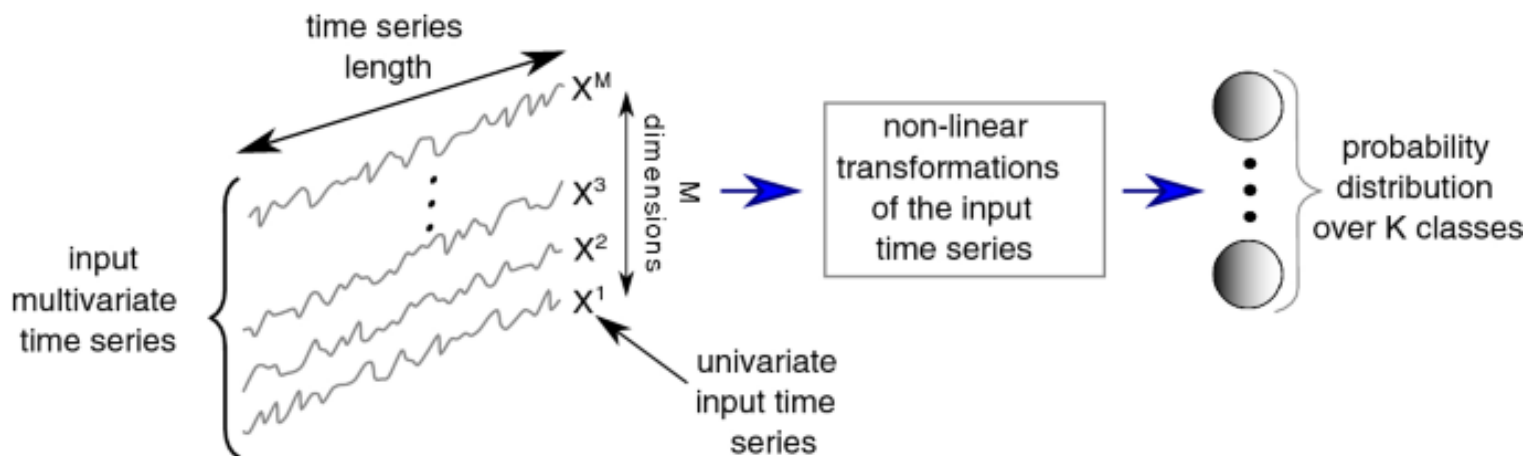


图 3: 针对时间序列数据分类的通用深度学习框架 [2]

2.1 基于统计特征的机器学习模型

根据文献 [3] 中所构建的 561 个基于时域特征及频域特征的变量，包括基于多种滤波器及快速傅里叶变换所构建的统计特征。在此基础上，考虑使用 K 近邻、朴素贝叶斯、决策树、支持向量机和随机森林等潜在模型。

2.2 基于 LSTM 的深度学习模型

由于 LSTM 模型可以很好地捕捉时序数据中的信号，因此，在此处我们不做什么预处理，直接将数据放入模型之中。

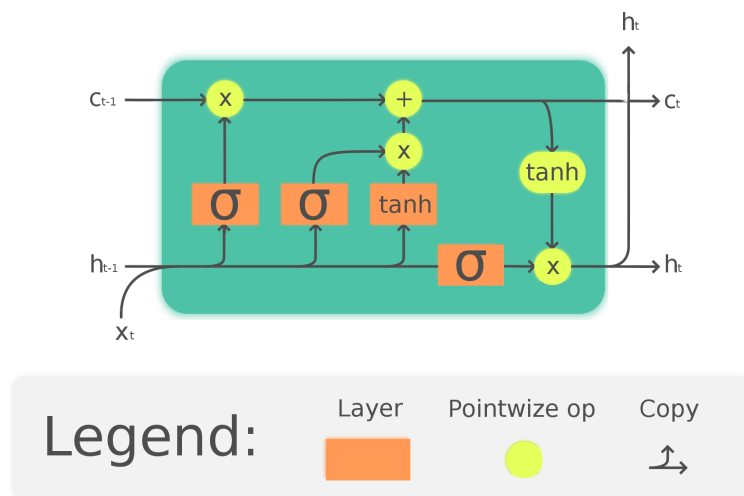


图 4: LSTM (Long short-term memory) 模型

注 "Pytorch-Lightning" 重构你的 PyTorch 代码，抽出复杂重复部分，让你专注于核心的构建，让你的实验更快速更便捷地开展迭代。参考示例：[SignorinoY/immortal-python](#)

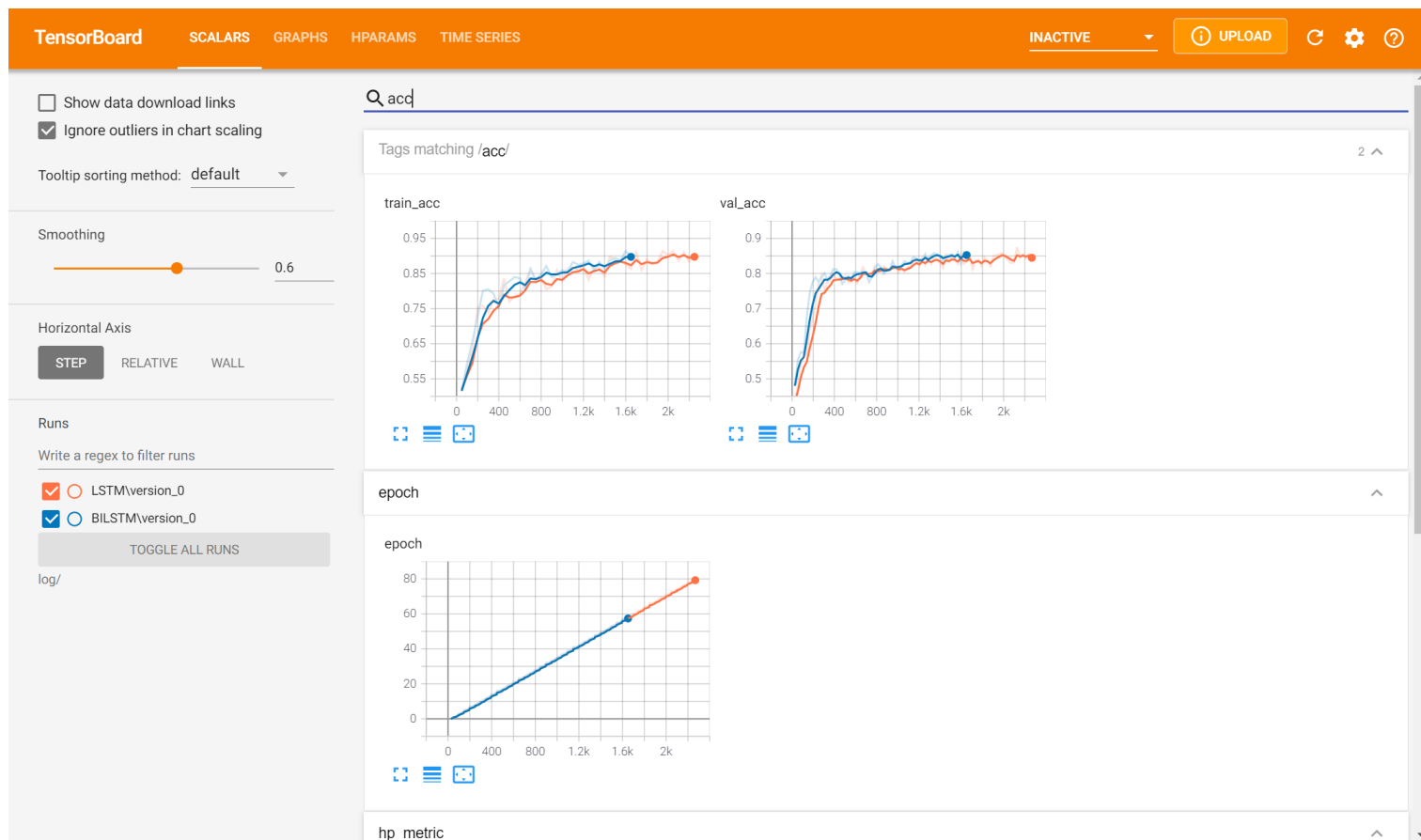


图 5: TensorBoard 实时运行截图

3 模型评价

在均使用模型默认参数的情况下，基于同一训练集（样本量：7352）进行训练，并在同一测试集（样本量：2947）上进行预测，评价模型的预测效果。

模型	准确率	Macro F1	模型	准确率	Macro F1
K 近邻	90.16%	0.90	随机森林	92.60%	0.93
朴素贝叶斯	77.02%	0.77	LSTM	71.50%	0.71
决策树	85.65%	0.86	BiLSTM	73.29%	0.72
支持向量机	95.05%	0.95			

表 1: 同一测试集下, 不同模型的预测结果

此项目代码在 [SignorinoY/corruptor](#) 处可见。

参考文献

- [1] ANGUIA D, GHIO A, ONETO L, et al. A public domain dataset for human activity recognition using smartphones[C/OL]// 2013: 437-442[2021-06-06]. <https://upcommons.upc.edu/handle/2117/20897>.
- [2] ISMAIL FAWAZ H, FORESTIER G, WEBER J, et al. Deep learning for time series classification: a review[J/OL]. Data Mining and Knowledge Discovery, 2019, 33(4): 917-963[2021-06-06]. <https://doi.org/10.1007/s10618-019-00619-1>.
- [3] ANGUIA D, GHIO A, ONETO L, et al. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine[C/OL]//Springer, 2012: 216-223[2021-06-06]. <https://upcommons.upc.edu/handle/2117/101769>. DOI: [10.1007/978-3-642-35395-6_30](https://doi.org/10.1007/978-3-642-35395-6_30).