

机器学习预测工科本科生英语学习的自主性

北京航空航天大学电子信息工程学院 吴晓君* 熊虹婷 王向荣 杨顺川 谢晋东 张有光

【摘要】在2020年特殊时期，大中小学的学生们均在家进行自主学习。这个过程中，良好的学习习惯和自律程度对学习效果有极大的影响。笔者通过为工科大三和四本科学生开设的《科技文献导读和科技论文写作》课程中增加每日英语听读打卡学习的方式，收集了学生在2020年3月和4月两个月学生学习打卡情况，采用机器学习的方法预测学生自觉学习的可能性。该研究发现，老师的督促、学生的学习兴趣、学生的自律程度对自主学习影响极为明显。

【Abstract】In the special period of 2020, students in elementary school, middle schools and universities have to stay at home and implement self-directed studies. In this process, good study habits and self-discipline have a great influence on the learning effect. In the “Literature Review and Scientific Paper Writing” course, the authors requested junior and senior students to practice listening and reading English every day, and analyzed the learning habit in the March and April of 2020. Then, they used machine learning method to predict the possibility of students conscious learning. It is found that teacher supervision, English learning interest and self-discipline of students have the most significant impacts on self-regulated learning process.

【关键词】自律；机器学习；英语；工科本科生

【Keywords】Self-discipline; machine learning; English; engineering undergraduates

科技英语对培养工科学生国际化水平，提升学校国际影响力等方面具有重要的作用。2020年春季开设的《文献导读和科技论文写作》课程，是在特殊时期（抗疫期间）采取的特殊方式（腾讯会议），并以自主学习为主的一门本科生选修课程。为了全面提升工科学生在英语听说读写方面的综合能力，在将来的学习和工作中能够更好地阅读英文资料，参与英文交流，发表英文演讲，善于撰写英文科技论文等，本课程提供了丰富而多样化的学习方式，主要包括：（1）每日听读打卡，训练学生英语综合听读能力；（2）课程讲解和自主学习，让学生明白科技文献阅读和科技论文基本写作方法；（3）在线模拟国际会议口头报告，让学生感受国际会议的气氛和进行英文口头演讲的尝试；（4）旁听研究生组会和学术报告，激发本科生科学研究的兴趣，拓展学生学术视野，了解科学研究的基本模式；（5）每月学习感想总结，及时了解学生的学习状况，做到查漏补缺。

在以上的多样化课程学习内容中，每日英语听读打卡学习可让老师每日都能与学生有所沟通和交流，让学生感觉老师就在身边，更加亲切。这样的方式也方便老师及时掌握学生的学习状况，有问题尽早解决。因此，这项内容在这门课程的教学过程中起到了非常重要的作用。学生通过阅读老师精心挑选的乐观积极、激励人心的英语散文，在疫情期间不仅可以从心灵上受到鼓舞，而且还能让学生们在家期间忙起来、学起来，为将来进一步深造或就业做好准备。从这样的学习中，我们也能部分地看出现在大学生，尤其是工科学生对英语学习的态度，以及坚持学习的可能性。

在疫情期间推行的自主学习的教学活动中，自律起到了至关重要的作用。由于受到客观环境的影响，以前的教室上课，变成了现在在家自主学习。没有老师的当

面督促，没有宿舍同学的互相激励，好习惯和自律则在自助学习中显得尤为重要。本文针对2020年春季学期大三和四学生同上的一门《文献导读和科技论文写作》课程进行了工科本科生科技英语学习自律性方面的研究。选修这门课程的学生共53人，于2020年2月28日要求学生每日进行英语听读打卡，共有53名学生参与打卡，其中大四（16个2016级，5个2015级，1个2014级）学生22名，大三（2017级）学生31名参与打卡。现将53名学生英语听读打卡分析如下：

一、按照打卡的数据进行听读遍数整理排名

分析打卡数据可知，学生每天听英语的遍数高于读英语的遍数，81.82%的大四学生和87.10%的大三学生能保证每天听英语两遍以上；59.09%的大四学生和74.19%的大三学生能保证每天读英语两遍以上。每日练习的英语材料为新东方英语《生而为赢》美文系列。该系列每篇文章阅读时间平均约为3分钟，则大部分学生可以保证每天15分钟的英语学习时间。这对养成英语学习习惯，提高口语和听力能力具有很大的帮助。共有27人全程坚持（一天没落下）每天英语打卡学习，其中大四学生10人，大三学生17人，分别占大四和大三学生打卡总人数的45.45%和54.84%。

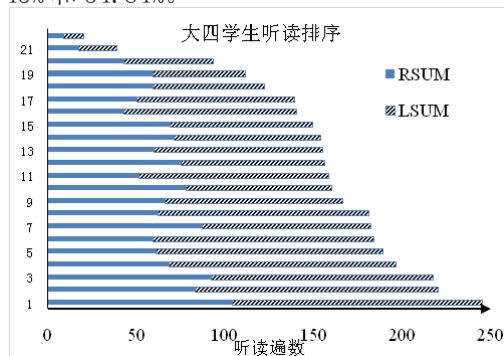


图1 大四听读遍数总和排序图

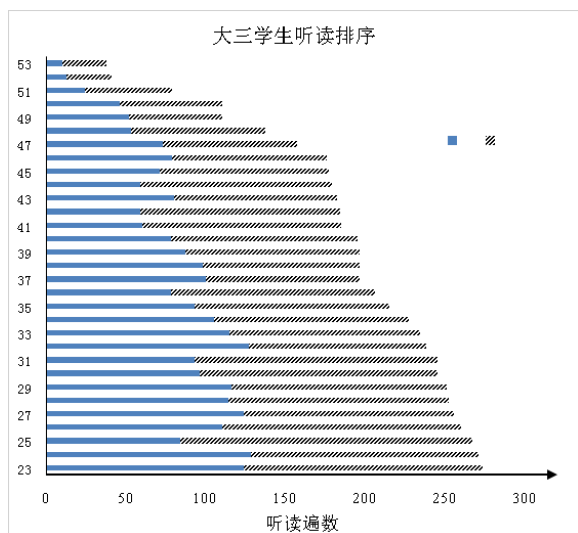


图2 大三听读遍数总和排序图

二、坚持程度分析

在打卡初期听读遍数较少,第6天达到听读遍数最大值,之后遍数开始下降,在第14天和第18天听读遍数较多,在第20天、第25天、第28天听读遍数较少。通过查阅日历发现听读遍数和是否是节假日无关,因此查阅了上述天数听读的英文文章,发现第14天和第18天的英文文章较长且存在较多生词,而第20天、25天和28天的英文文章生词较少。由此可以推断,英文文章的难易程度是影响学生英语听读遍数的主要因素。为了减少该因素对英语听读遍数的影响,将每2天和每3天的听读平均值求和发现,每日读英语遍数比听英语遍数先趋于平衡。在英语打卡活动进行到第9天时,每日听英语遍数和读英语遍数均达到的最大值,此后听英语遍数小幅度下降,听英语遍数逐渐稳定在每天4遍,读英语遍数稳定在每天3遍。

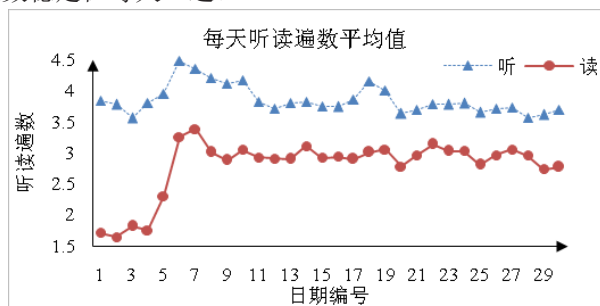


图3 每天听读遍数平均值

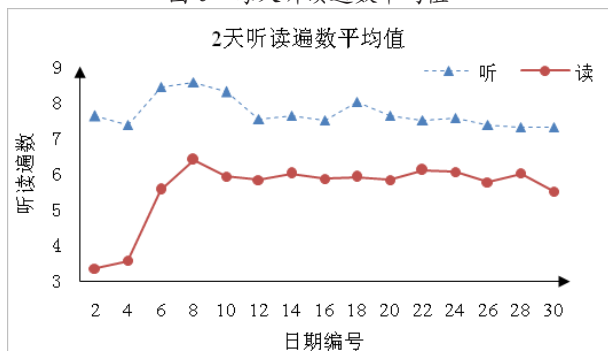


图4 每2天听读遍数平均值

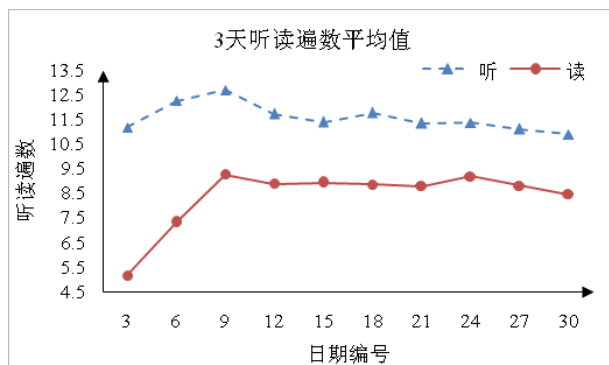


图5 每3天听读遍数平均值

三、对听读遍数产生影响的因素

在发现英文文章的难易程度对听读遍数影响规律之后,我们又对年级、生源地、六级通过情况、以及英语学习兴趣对每日听读遍数的影响进行了分析。

1. 年级影响

分别统计大四和大三学生听读遍数的区别可知,大四学生听英语遍数略低于大三学生,而读英语遍数则高于大三学生,可见大四学生对于英语口语能力的需求大于大三学生。这可能与他们即将毕业,面临考研英语口语复试、出国以及找工作的需求有一定关系。

2. 生源地

我国国土辽阔,东南沿海地区较西北内陆地区发展快速,这势必导致教育资源的不公平。在学生进入到大学学习前他们所接受的英语学习方法不同,积累的英语学习经验也不同,所处的英语学习环境也有所不同。所以,将英语打卡学生生源根据我国四大经济区划分规则进行分类,统计他们的英语听读次数可知,无论是听英语遍数还是读英语遍数,东部地区均高于中部和西部地区。这个很好的反应了进入大学前的学习经历在一定程度上会影响到学生进入大学的学习习惯。由于东北地区学生人数只有4人,则没有很好的反应这个相关关系。

东部地区、中部地区和西部地区学生听英语遍数基本没有很大区别,但是东部地区学生读英语遍数则明显高于中部地区和西部地区学生,其原因可能是东部地区学生对于自己口语能力要求更高一些。

3. 六级通过情况

此次参与英语打卡的53名同学全部通过英语4级考试,有一半的学生通过了英语6级考试。将学生按照是否通过6级进行分类绘制英语听读遍数可知,是否通过6级考试对英语听读遍数影响较大。其中,未通过6级考试的同学听英语遍数高于通过6级考试的同学,读英语遍数低于通过6级考试的同学。由此可见,每位学生在进行听读练习时会根据个人需求调整听读遍数,未通过6级考试的学生会加大听英语的力度,增加自己的英语听力能力;通过6级考试的学生会加大读英语的力度,增加自己的英语口语交流能力。因此,从数据上反应,学

生们认为每天坚持英语听读学习对于提高英语6级通过率可能有一定帮助。

4. 学习兴趣

兴趣是学习的最大动力,人们总是会对自己感兴趣的内容投入百分百的努力,并且乐在其中十分享受学习的过程。按照学生是否喜欢英语学习进行分类可得,只有少部分学生有学习英语的兴趣。因此,他们每日听读英语的遍数明显高于英语学习兴趣一般或较低的同学,并且更能坚持每天进行英语学习打卡。无论是听英语遍数还是读英语遍数,具备英语学习兴趣的学生打卡平均值均高于不具备英语学习兴趣的学生10遍以上。因此,想要更好的学习英语需要找到学习英语的乐趣,使学习英语变成一种很快乐的过程。

四、大三和大四同学的听读遍数区别

我们还分别统计了大四和大三学生听读英语遍数,发现大四学生和大三学生听英语遍数均高于读英语遍数,而大四学生中只有1人听英语遍数少于读英语遍数,大三学生中有2人听英语遍数少于读英语遍数。大四学生听英语遍数平均值比大三学生少20遍,读英语遍数少21遍。大四的学生即将毕业,他们面临这就业压力、考研压力,可以用来学习英语的时间明显少于大三的学生。

五、预测模型

通过上文分析可知,一个学生英语学习经历、是否通过6级考试、是否具备英语学习兴趣以及是否是毕业生都会在一定程度上影响他们英语学习坚持的可能性。将48名学生(总数为53人,其中有5名学生未参与模型信息统计)随机分为训练集和测试集,其中训练集38名学生,测试集10名学生,采用ECOC-SVM模型。该模型适用于小样本训练,并且SVM在二分类中具有较好的分类效果。并且采用ECOC纠错编码,可进一步提高分类准确率。该模型利用线性核函数进行训练,10名测试集学生中,7人预测准确,3人预测错误,测试集准确率为70%。

该模型训练集准确率为84%,误差来源于大四和大三学生、生源地、六级通过情况等属于人工挑选的特征向量,本身与是否能坚持学习的相关性可能不是很大,所以模型训练误差也大,导致测试集准确率不是非常高。要想获得更高的预测准确率,可再扩大训练集数量,或者挑选更为合适的特征向量(可以反应学生坚持学英语的因素)。进一步优化模型参数,比如将线性核函数更改为径向基函数或高斯函数,也有可能提高预测准确率。

六、模型应用

通过三月份的数据得到预测模型,将其应用到四月份打卡数据中。截止四月份统计时间共十天,鉴于部分同学中期答辩,后期会补充上打卡,因此在十天中坚持

八天则认为该同学可以坚持英语学习。模型预测的48名同学有28名同学可能会在四月份继续坚持学习,但是实际只有19名学生坚持英语学习,模型预测准确率为60.42%。

为验证30天的学习打卡是否可以养成英语学习的习惯,在四月份学习打卡过程中,老师故意通过发电子奖状来激励学生,也没有通过每日将打卡记录表定期发送到群里来督促学生学习,全凭学生自觉进行英语学习。从这十天的实测数据发现,有20名同学坚持了打卡,并且有13名学生没有一天参与英语学习打卡活动中,而在三月份有27名同学坚持了打卡,所有学生均参与了英语学习打卡。由此可知,有效的激励机制和老师的督促对于无法自己坚持学习的学生是很有必要的,而对于本身具备良好的学习习惯和有极高自律性的学生而言影响则较低。

七、总结与展望

收集了2020年3月和4月疫情期间开设的《文献导读与科技论文写作》课程学生每日英语听读学习打卡数据,通过机器学习的模型对大三和大四学生在这门课程中影响英语学习的可能因素以及坚持程度进行了分析和预测。实验结果证实,老师的督促、学习的兴趣、文章的内容、学生的年级等因素对英语学习有极大的影响,六级通过情况和生源地等影响不是非常明显。三月份预测结果与四月份实验结果相差甚远再一次证实,老师督促、严厉程度、激励机制起着至关重要的作用。该实验结果再次表明,28天的确只能形成一个模式,为学生们养成学习的好习惯,老师的督促可能需要更长的时间!

【参考文献】

- [1] 师莹. 小组主题汇报提升专门用途英语学生参与度的行动研究——以科技英语为例[J]. 教书育人(高教论坛), 2019(27):108-109.
- [2] 吴丽萍. 非英语专业大学新生英语学习动机情况调查及其与学习者自主的相关性研究[J]. 海外英语, 2020(04):142-145.
- [3] 江畅. 高科技时代的人类自律及其构建[J]. 湖北大学学报(哲学社会科学版), 2020, 47(02):1-10+167.

【基金项目】北京航空航天大学2018年教学改革项目“探索适合本科生参与科研的方法与实践研究”。

【通讯作者】吴晓君(1983—),女,汉族,重庆人,博士,副教授,研究方向:太赫兹科学与技术