

分类号: _____

单位代码: 10335

学 号: 21621196

浙江大学

硕士学位论文



论文题目: 基于分类模型的英语
语法纠错算法研究

作者姓名: 诸凯丽

指导教师: 蔡登

学科 (专业): 计算机科学与技术

研究方向: 自然语言处理

所在学院: 计算机科学与技术学院

提交日期 二〇一八年十二月十日

A Dissertation Submitted to Zhejiang
University for the Degree of
Master



**TITLE: Research on Classification Model of
English Grammar Error Correction**

Author: Kaili Zhu

Supervisor: Deng Cai

Subject: Computer Science and Technology

Research Domain: Natural Language Processing

College: College of Computer Science and Technology

Submitted Date 2018-12-10

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 浙江大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：

签字日期： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解 浙江大学 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权浙江大学可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：

导师签名：

签字日期： 年 月 日 签字日期： 年 月 日

摘要

这是一篇有关法器钉耙的研究。以往法器研究都是从西方视角出发，没有考虑东方文化背景的影响。本研究结合钉耙在农业中的广泛应用，对钉耙的作用进行开创性研究。

扯几句，再扯几句。

关键词：钉耙，成本，法器， \LaTeX

Abstract

Rake is good. My rake has nine prongs. Rake can be used in farm. Blur, blur, blur...

Keywords: rake, cost, X_YTeX

目录

摘要 I

Abstract..... II

第 1 章 绪论..... 1

 1.1 字体示例..... 1

 1.2 各种示例..... 2

 1.2.1 图片示例 2

 1.2.2 表格示例 4

 1.2.3 公式示例 4

 1.2.4 定理示例 5

 1.2.5 算法示例 6

 1.2.6 参考文献 6

第 2 章 第二章..... 7

参考文献..... 8

攻读硕士学位期间的主要研究成果 9

致谢 10

插图

1.1 玉照, 玉做的照片, 通常会令人联想到浴照。 3

1.2 你挑着担, 我写代码... 3

表格

1.1 模板中的表格宏包 4

第 1 章 绪论

这是的示例文档。格式样列很大一部分是直接从薛瑞尼的《清华大学学位论文 L^AT_EX 模板》，<http://thuthesis.sourceforge.net/> 示例文档中拿来的。

1.1 字体示例

九齿钉耙学名上宝沁晶耙，是俺的武器。九齿钉耙并非普通的农具，而是由太上老君¹用神冰铁亲自锤炼，借五方五帝、六丁六甲之力锻造而成，有诗为证：

字体	示例
fangsong	老君自己动铃锤，荧惑亲身添炭屑。 五方五帝用心机，六丁六甲费周折。 造成九齿玉垂牙，铸就双环金坠叶。
heiti	身妆六曜排五星，体按四时依八节。 短长上下定乾坤，左右阴阳分日月。 六爻神将按天条，八卦星辰依斗列。
kaishu	名为上宝沁金耙，进与玉皇镇丹阙。 因我修成大罗仙，为吾养就长生客。 敕封元帅号天蓬，钦赐钉耙为御节。
lishu	举起烈焰并毫光，落下猛风飘瑞雪。 天曹神将尽皆惊，地府阎罗心胆怯。 人间那有这般兵，世上更无此等铁。
songti	随身变化可心怀，任意翻腾依口诀。 相携数载未曾离，伴我几年无日别。 日食三餐并不丢，夜眠一宿浑无撇。
youyuan	也曾佩去赴蟠桃，也曾带他朝帝阙。 皆因仗酒却行凶，只为倚强便撒泼。

¹太上老君，三清之第三位。又称“道德天尊”、“混元老君”、“降生天尊”、“太清大帝”等。

上天贬我降凡尘，下世尽我作罪孽。
default 石洞心邪曾吃人，高庄情喜婚姻结。
这钯下海掀翻龙鬲窝，上山抓碎虎狼穴。
诸般兵刃且休题，惟有吾当钯最切。
相持取胜有何难，赌斗求功不用说。
何怕你铜头铁脑一身钢，钯到魂消神气泄！

1.2 各种示例

没有规定英文字体，将中、英文统一设置为仿宋。

当存在连续多个表格、图片时， \LaTeX 的排版效果不佳，所以这里不时插播些王勃的《滕王阁序》，有空的话也看两眼。

《滕王阁序》第一段：豫章故郡，洪都新府。星分翼轸，地接衡庐。襟三江而带五湖，控蛮荆而引瓯越。物华天宝，龙光射牛斗之墟；人杰地灵，徐孺下陈蕃之榻。雄州雾列，俊采星驰。台隍枕夷夏之交，宾主尽东南之美。都督阎公之雅望，棨戟遥临；宇文新州之懿范，檐帷暂驻。十旬休暇，胜友如云；千里逢迎，高朋满座。腾蛟起凤，孟学士之词宗；紫电青霜，王将军之武库。家君作宰，路出名区，童子何知，躬逢胜饯。

《滕王阁序》第二段：时维九月，序属三秋。潦水尽而寒潭清，烟光凝而暮山紫。俨骖騑于上路，访风景于崇阿。临帝子之长洲，得天人之旧馆。层台耸翠，上出重霄；飞阁翔丹，下临无地。鹤汀凫渚，穷岛屿之萦回；桂殿兰宫，即冈峦之体势。披绣闼，俯雕甍：山原旷其盈视，川泽纡其骇瞩。闾阎扑地，钟鸣鼎食之家；舸舰迷津，青雀黄龙之轴。云销雨霁，彩彻区明。落霞与孤鹜齐飞，秋水共长天一色。渔舟唱晚，响穷彭蠡之滨；雁阵惊寒，声断衡阳之浦。

1.2.1 图片示例

附上照片一张，来源 Wikipedia。

我来插两张并排的图像，其实 \LaTeX 支持 jpg, png, 以及 eps 格式的图片，图1.2a是我的师傅，帅吧！



图 1.1 玉照, 玉做的照片, 通常会令人联想到浴照。



(a) 师傅



(b) 大师兄

图 1.2 你挑着担，我写代码...

1.2.2 表格示例

L^AT_EX 中的表格不是很好搞，看起来很不直观，不过我们一般不会用太高端的表格，更多关于表格的请看 Helin Gai 的 L^AT_EX Manual^[1] 以及 Alpha Huang 的 lnotes^[2]

先看一个简单表格 1.1:

表 1.1 模板中的表格宏包

宏包	描述
longtable	绘制跨页的表格。
booktabs	三线表中的那三条线的命令来自这里。
caption2	用于设置标题很方便，已经 obsolated，不过 T _E XLive 中还有。
multirow	跨行的单元格用这个宏包。
dcolumn	想让表格小数点对齐吗？用这个宏包吧。
colortbl	表格上色。自己看着爽而已，打印出来都是黑白的。
threeparttable	用来给表格添加脚注啥的很方便。
array	忘了用来做什么了，但似乎很重要。

1.2.3 公式示例

我公式用非常少，所以公式的情况并不了解，这部分示例也完全照搬。如果有什么问题或要求，可以贴到 88 T_EX 版。

贝叶斯公式如式 (1.1)，其中 $p(y|\mathbf{x})$ 为后验； $p(\mathbf{x})$ 为先验；分母 $p(\mathbf{x})$ 为归一化因子。

$$p(y|\mathbf{x}) = \frac{p(\mathbf{x}, y)}{p(\mathbf{x})} = \frac{p(\mathbf{x}|y)p(y)}{p(\mathbf{x})} \quad (1.1)$$

论文里面公式越多，T_EX 就越 happy。再看一个 amsmath 的例子：

$$\det \mathbf{K}(t = 1, t_1, \dots, t_n) = \sum_{I \in \mathbf{n}} (-1)^{|I|} \prod_{i \in I} t_i \prod_{j \in I} (D_j + \lambda_j t_j) \det \mathbf{A}^{(\lambda)}(\bar{I}|\bar{I}) = 0. \quad (1.2)$$

大家在写公式的时候一定要好好看 amsmath 的文档，并参考模板中的用法：

$$\begin{aligned}
& \int_a^b \left\{ \int_a^b [f(x)^2 g(y)^2 + f(y)^2 g(x)^2] - 2f(x)g(x)f(y)g(y) dx \right\} dy \\
&= \int_a^b \left\{ g(y)^2 \int_a^b f^2 + f(y)^2 \int_a^b g^2 - 2f(y)g(y) \int_a^b fg \right\} dy \quad ([b])
\end{aligned}$$

其实还可以看看这个多级规划：

$$\left\{ \begin{array}{l} \max_x F(x, y_1^*, y_2^*, \dots, y_m^*) \\ \text{subject to:} \\ G(x) \leq 0 \\ (y_1^*, y_2^*, \dots, y_m^*) \text{ solves problems } (i = 1, 2, \dots, m) \\ \left\{ \begin{array}{l} \max_{y_i} f_i(x, y_1, y_2, \dots, y_m) \\ \text{subject to:} \\ g_i(x, y_1, y_2, \dots, y_m) \leq 0. \end{array} \right. \end{array} \right. \quad (1.3)$$

1.2.4 定理示例

THUThesis 定义了很多定理格式，THUThesis 下面都是从来的例子。我自己只用到了“假设”，对于公式、证明之类的格式完全无知，你如有其他需要可以在 88 T_EX 版上提出来，我再添加。

假设 1.1：待月西厢下，迎风户半开；隔墙花影动，疑是玉人来。

$$c = a^2 - b^2 \quad (1.4)$$

$$= (a + b)(a - b) \quad (1.5)$$

定义 1.1：子曰：「道千乘之国，敬事而信，节用而爱人，使民以时。」

定理 1.1：犯我强汉者，虽远必诛。

——陈汤（汉）

证明 1.1: 天不言自高, 水不言自流。

$$\begin{aligned}\varphi(x, z) &= z - \gamma_{10}x - \gamma_{mn}x^m z^n \\ &= z - Mr^{-1}x - Mr^{-(m+n)}x^m z^n\end{aligned}$$

$$\zeta^0 = (\xi^0)^2, \quad (1.6)$$

$$\zeta^1 = \xi^0 \xi^1, \quad (1.7)$$

$$\zeta^2 = (\xi^1)^2, \quad (1.8)$$

1.2.5 算法示例

大家有时候还需要写一些算法的伪代码, 不用蛋疼, 模板里面已经有支持了, 看看算法1, 具体的实用还需要参考 CTAN 上的关于 `algorithmicx` 宏包的文档, 包括自定义关键字等等。

算法 1 把猪八戒放进冰箱

输入: N 头动物, 不知道是不是猪八戒

▷ 我是注释

输出: 你造吗, 这个程序木有输出, 想怎样?

1: 先热热身:)

2: **for** $i = 1, \dots, N$ **do**

▷ 我是个 For 循环

3: **if** 是猪八戒 **then**

4: 打开冰箱门, 把第 i 个放进冰箱, 关上冰箱门

5: **else**

6: 下一个

7: **end if**

8: **end for**

1.2.6 参考文献

参考文献建议使用 BiBTeX。

参考文献的例子: 书籍文献^[3;4]; 杂志文章, 单独的^[5], 并列的^[6;7], 合并的^[5-7]; 硕士论文^[8], 俺的旧文一篇; 博士论文^[9], 沙师弟的大作; 论文集或 Handbook^[10], 会议论文集^[11], 会议论文^[12]; 带页码脚标的^[3] 123。

第2章 第二章

这是第二章，目的是占个位置。你知道《浙大同学爱占坐》这首歌吗？

参考文献

- [1] Helin Gai. The Art of L^AT_EX[M]. Duke University, 2008.
- [2] Alpha Huang. L^AT_EX Notes[M]. 2008.
- [3] Donald E. Knuth. The T_EX Book[M]. Reading, MA: Addison-Wesley Publishing Company, 1989, 15th edition.
- [4] Michel Goossens, Frank Mittelbach, Alexander Samarin. The L^AT_EX Companion[M]. Reading, MA: Addison-Wesley Publishing Company, 1994.
- [5] P. S. Adler, B. Goldoftas, D. I. Levine. Flexibility versus efficiency? A case study of model changeovers in the Toyota production system[J]. Organization Science, 1999, 10:43–68.
- [6] Philippe Aghion, Jean Tirole. Formal and real authority in organizations[J]. Journal of Political Economy, 1997, 105(1):1–29.
- [7] D. Ahlstrom, G. D. Bruton. An institutional perspective on the role of culture in shaping strategic actions by technology-focused entrepreneurial firms in china[J]. Entrepreneurship: Theory & Practice, 2003, 26(4):52–69.
- [8] 猪八戒, 猪悟能, 猪刚鬣, 天蓬元帅. 论流体食物的持久保存 [D]. [硕士学位论文], 北京, 广寒宫大学, 2005.
- [9] 沙和尚. 论流沙河的综合治理 [D]. [博士学位论文], 北京, 流沙河大学, 2005.
- [10] Andreas Rauch. Characteristics of small-scale business owners: A meta-analysis on emergence and success. In W. D. Bygrave, C. G. Brush, M. Lerner, P. Davidsson, G. D. Meyer, J. Fiet, J. Sohl, P. G. Greene, A. Zacharakis, R. T. Harrison, eds., Frontiers of Entrepreneurship Research, Wellesley, Massachusetts: Babson College, 2003, 517–529.
- [11] Andreas Rauch. Characteristics of small-scale business owners: A meta-analysis on emergence and success[C]. In W. D. Bygrave, C. G. Brush, eds., Frontiers of Entrepreneurship Research. Wellesley, Massachusetts: Babson College, 2003, 517–529.
- [12] S. C. Selden. Testing a multi-dimensional model of organizational performance: Prospects and problems[C]. In Paper presented at the 7th National Public Management Research Conference. Washington, DC.: NPMRC, 2003, 222–234.

攻读硕士学位期间的主要研究成果

- 猪八戒, 猪悟能, 天蓬元帅, 等. 论流体食物的持久保存 [D]. 硕士学位论文. 北京: 广寒宫大学, 2005

致谢

感谢导师唐三藏的温柔和慈爱，不计时间给予俺无微不至地教诲，让俺在如雨般的唾沫星下滋润成长。

感谢师兄孙悟空的肢体教育，让俺成功减肥，俺和俺身上的每一块青紫都会记住你的。

感谢师弟沙悟净的任劳任怨，得失无争。

感谢高小姐的一片真情，可叹俺已遁入空门、修成正果。当年你在情人坡娇怨俺没貌没地位，现在俺终于位列仙班，可却注定要永远分离，缘何俺们总是“错位”……

猪悟能

2013 年 1 月于大雷音寺