

Auto Summarization Technology

Minchiuan Gao

minchiuan.gao@gmail.com

Outline

- i. Background;
- ii. Related Categories;
- iii. TextRank;
- iv. Sentence Embedding Based (KNN, LDA);
- v. Templates Based;
- vi. Other Methods;

Background

- The needs for converting a longer text to a shorter one.

网易娱乐7月21日报道 林肯公园主唱查斯特·贝宁顿Chester Bennington于今天早上，在洛杉矶帕洛斯弗迪斯的一个私人庄园自缢身亡，年仅41岁。此消息已得到洛杉矶警方证实。

洛杉矶警方透露，Chester的家人正在外地度假，Chester独自在家，上吊地点是家里的二楼。一说是一名音乐公司工作人员来家里找他时发现了尸体，也有人称是佣人最早发现其死亡。

林肯公园另一位主唱麦克·信田确认了Chester Bennington自杀属实，并对此感到震惊和心痛，称稍后官方会发布声明。Chester昨天还在推特上转发了一条关于曼哈顿垃圾山的新闻。粉丝们纷纷在该推文下留言，不相信Chester已经走了。

外媒猜测，Chester选择在7月20日自杀的原因跟他极其要好的朋友、Soundgarden（声音花园）乐队以及Audioslave乐队主唱Chris Cornell有关，因为7月20日是Chris Cornell的诞辰。而Chris Cornell于今年5月17日上吊自杀，享年52岁。Chris去世后，Chester还为他写下悼文。

对于Chester的自杀，亲友表示震惊但不意外，因为Chester曾经透露过想自杀的念头，他曾表示自己童年时被虐待，导致他医生无法走出阴影，也导致他长期酗酒和嗑药来疗伤。目前，洛杉矶警方仍在调查Chester的死因。

据悉，Chester与毒品和酒精斗争多年，年幼时期曾被成年男子性侵，导致常有轻生念头。Chester生前有过2段婚姻，育有6个孩子。

林肯公园在今年五月发行了新专辑《多一丝曙光One More Light》，成为他们第五张登顶Billboard排行榜的专辑。而昨晚刚刚发布新单《Talking To Myself》MV。

林肯公园主唱查斯特·贝宁顿Chester Bennington，今天早上，在洛杉矶帕洛斯弗迪斯的一个私人庄园自缢身亡，年仅41岁。粉丝们纷纷在该推文下留言，不相信Chester已经走了。外媒猜测，Chester选择在7月20日自杀的原因跟他极其要好的朋友、Soundgarden乐队以及Audioslave乐队主唱Chris Cornell有关，因为7月20日是Chris Cornell的诞辰。今年5月17日上吊自杀，享年52岁，去世后，Chester还为他写下悼文。

Background

- Multiply documents merge to single document.



Background

- Auto report.



项目	本周	比上周增减 (%)
成交数量 (万股)	58,805.77	-42.85%
其中：做市方式	24,375.63	-44.29%
协议方式	34,430.14	-41.80%
成交金额 (万元)	334,563.93	-47.46%
其中：做市方式	138,318.11	-43.83%
协议方式	196,245.82	-49.76%
成交均价 (元)	5.69	-8.07%
其中：做市方式	5.67	0.82%
协议方式	4.70	-13.67%

工作周报
姓名：小龙女 部门：研发一部 职位：经理助理 第16周
周一：整理上周各项目进展报告，向经理汇报汇总结果 周二：电话拜访客户，收集并汇总客户在产品使用过程中的问题 周三：联系企业内训讲师，初审内训大纲，安排内训时间、地点等 周四：整理研发项目的文档，主持并参与项目经理会议 周五：收集各项目组进展报告，配合经理抽查项目完成情况 周六：加班，协助组织并参与企业内训 周日：休息
遇到的问题： 一切顺利！
收获和体会： 这周好忙，周一至周五都加班！

工作周报
姓名：小龙女 部门：研发一部 职位：经理助理 第17周
周一：整理上周各项目进展报告，向经理汇报汇总结果 周二：电话拜访客户，收集并汇总客户在产品使用过程中的问题 周三：与经理一起参加公司中高层干部会议，讨论公司规划 周四：整理研发项目的文档，主持并参与项目经理会议 周五：收集各项目组进展报告，配合经理抽查项目完成情况 周六：休息 周日：休息
遇到的问题： 一切顺利！
收获和体会： 这周好忙，周一至周五都加班！

Auto Summarization

- Complexity
 - Concise
 - Fluently
 - Semantic
- Diversity
 - Document Format
 - Target(report, news, table, etc)
 - Topic Analysis
 - LDA
 - Sentence Score
 - Etc



Abstractive

导语

近日，一则高考改革方案引起了舆论哗然，因为其中提到“探索开展综合评价录取模式改革试点，高考成绩占比不低于60%”。有网民表示，以后高考不仅要拼分，还要拼爹。其实，浙江、上海、山东等地在前几年已经开始了综合评价录取的试点，高考“一考定终身”的局面在不远的将来可能真的会成为历史，这到底是好事还是坏事？

要 • 点 • 速 • 读

- 1 综合评价录取正在悄然改变高考格局，靠裸分上名校会越来越难。
- 2 相关研究数据表明，综合评价录取更有利于城镇学生脱颖而出。
- 3 高考兼具教育因素和社会因素，其改革必须固守公平。

▷ 什么是综合评价录取？



Extractive

网易娱乐7月21日报道 林肯公园主唱查斯特·贝宁顿Chester Bennington于今天早上，在洛杉矶帕洛斯弗迪斯的一个私人庄园自缢身亡，年仅41岁。此消息已得到洛杉矶警方证实。

洛杉矶警方透露，Chester的家人正在外地度假，Chester独自在家，上吊地点是家里的二楼。一说是一名音乐公司工作人员来家里找他时发现了尸体，也有人称是佣人最早发现其死亡。

林肯公园另一位主唱麦克·信田确认了Chester Bennington自杀属实，并对此感到震惊和心痛，称稍后官方会发布声明。Chester昨天还在推特上转发了一条关于曼哈顿垃圾山的新闻。粉丝们纷纷在该推文下留言，不相信Chester已经走了。

外媒猜测，Chester选择在7月20日自杀的原因跟他极其要好的朋友、Soundgarden（声音花园）乐队以及Audioslave乐队主唱Chris Cornell有关，因为7月20日是Chris Cornell的诞辰。而Chris Cornell 于今年5月17日上吊自杀，享年52岁。Chris去世后，Chester还为他写下悼文。

对于Chester的自杀，亲友表示震惊但不意外，因为Chester曾经透露过想自杀的念头，他曾表示自己童年时被虐待，导致他医生无法走出阴影，也导致他长期酗酒和嗑药来疗伤。目前，洛杉矶警方仍在调查Chester的死因。

据悉，Chester与毒品和酒精斗争多年，年幼时期曾被成年男子性侵，导致常有轻生念头。Chester生前有过2段婚姻，育有6个孩子。

林肯公园在今年五月发行了新专辑《多一丝曙光One More Light》，成为他们第五张登顶Billboard排行榜的专辑。而昨晚刚刚发布新单《Talking To Myself》MV。

林肯公园主唱查斯特·贝宁顿Chester Bennington，今天早上，在洛杉矶帕洛斯弗迪斯的一个私人庄园自缢身亡，年仅41岁。粉丝们纷纷在该推文下留言，不相信Chester已经走了。外媒猜测，Chester选择在7月20日自杀的原因跟他极其要好的朋友、Soundgarden乐队以及Audioslave乐队主唱Chris Cornell有关，因为7月20日是Chris Cornell的诞辰。今年5月17日上吊自杀，享年52岁，去世后，Chester还为他写下悼文。

Programming Example

- 1. TextRank
- 2. Sentence Embedded
- 3. Template Based

Assignments

- Implementing the auto-summarization from dataset.
- Reading the paper *Text Summarization Techniques: A Brief Survey*