# 深度学习与时间嵌入预测情绪

概述

Emoji的使用已经成为一种新的社交方式，这一点很重要，因为它有助于改进聊天应用等通信系统。本文通过研究表情符号的用法和语义来分析表情符号用法的季节性变化。此外，他们还开发了一个基于时间信息的情绪预测模型。

贡献

过去曾进行过多项情绪预测研究（见著名著作），但没有考虑到时间信息。探讨了表情符号与季节性事件之间的时间相关性，并将其用于消除表情符号的歧义。

例子

以叶三叶草表情符号（leaf clover emoji）为例，它通常与全年的好运祝福联系在一起，除了在3月份，它主要用于表达与聚会和饮酒有关的事件情况（由于圣帕特里克节）。

挑战

* 这项研究表明，时间信息对于情绪预测是有用的，即使是对于与时间无关的情绪（💪和❤65039;）。
* 表情符号天生就是主观的，这就是为什么很难分析它们的语义。

数据集

Twitter用于收集1亿条美国tweets的语料，其组织如下：

* 季节表情数据集 - 数据按季节分为四个子集：春、夏、秋和冬（见下图）
* Emoji预测数据集 - 数据缩减为只包含一个频繁Emoji的tweets（Emoji必须属于300个频繁Emoji）

季节性表情符号的语义与用法

利用季节性数据集的四个子集训练跳格词嵌入模型。这些模型提供的信息基本上有助于根据表情符号之间的语义相似性来描述表情符号。（详见论文）

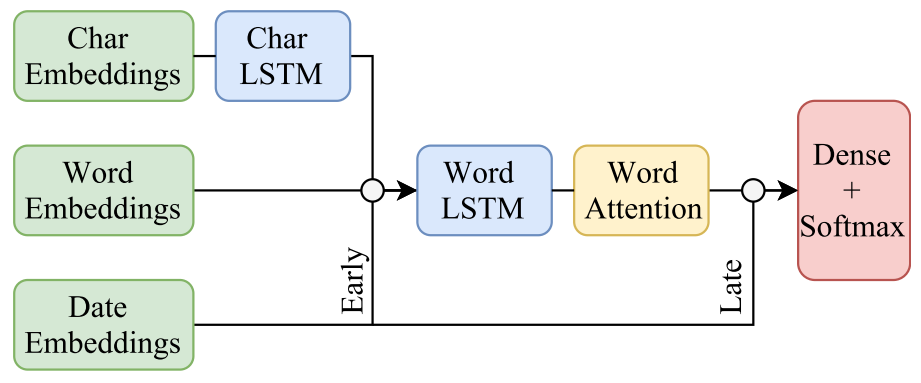
通过比较嵌入空间中与每个表情符号相关联的前10个表情符号，发现与音乐、动物、糖果和情感相关的表情符号不受季节性的影响（例如，🎶，🎼，🍦，🐠，😂，🎸）。这意味着这些表情符号在不同季节都保留了意义。

相比之下，与运动相关的表情符号（如🏀、🏆）在不同季节的含义有所不同，这可能是因为运动的旺季。另一个与学校有关的有趣的表情符号（🎓），随着季节的变化而变化；在春天，它与聚会表情符号有关；在秋天，它与学校相关的表情符号有关。看看下面的图中的松树表情符号（🌲）的前十个相关表情符号，你不认为这是季节性的吗？你能猜出原因吗？（提示：户外vs圣诞节）。（大量有趣的发现见论文）



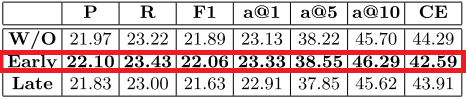
情绪预测

第二个数据集包括300个emoji类和90万条tweets（每个类3000条tweets），用于emoji预测。表情预测模型的体系结构如下：通过早期融合方法和后期融合方法将字符嵌入、单词嵌入和数据嵌入结合起来。这产生了两种模式（早期和晚期）。第三个模型是完全忽略数据嵌入的训练模型（W/O）。（有关如何构造这些嵌入的信息，请参阅论文）

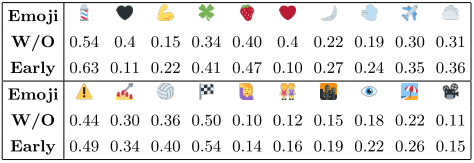


结果

下表列出了所有车型的精确性、召回率和F1分数。我们可以观察到，通过使用早期融合来结合时间信息，早期模型的性能优于其他模型。



下表列出了在F1得分上有更高收益的表情符号（无日期和提前日期）。您可以肯定地看到，许多emojis都是季节性的（例如，🍀，🌒），因此可以从日期嵌入中获益。即使与时间无关的表情符号（例如，🖤，❤️，💪）也从时间信息中受益。



结论与未来工作

* 提出了一种基于深度神经网络的多模式情感预测体系结构。
* 对一天或一周中特定时间段的emoji语义和用法进行更多的分析可能有助于改进数据嵌入和整体预测模型。
* This work has a lot of room for improvement and it could be a very interesting topic to combine with emotion recognition, event detection, and computational health studies.

**References**

* Ref: <https://arxiv.org/abs/1805.00731> — “Exploring Emoji Usage and Prediction Through a Temporal Variation Lens**”**