## 2.产业发展模块

互联网大数据对资源能源上下游产业的发展情况也能起到很好的监测预测作用。在不同产业的产品需求、价格和销售情况，以及产业的发展景气水平监测预测上都已经积累了一定的研究内容。主要利用互联网大数据所反映的人们在现实世界里的需求、兴趣、和关注，来和产业相关的社会经济活动联系起来。

2.1产品需求的监测预测

2.2产品价格的监测预测

2.3产品销售情况的监测预测

大数据在多种产品的销售情况监测预测上都被验证具有良好的作用。以下是部分代表性研究在不同产品领域上的总结。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 相关研究 | 研究内容 | 数据源 | 方法 | 效果 |
| Goel S,Hofman J M,Lahaie S,et al（2010） | 电影票房收入和游戏销售变化 | 2008-2009Yahoo!网页搜索日志与119部电影的票房数据（互联网电影数据库IMDB），选择与IMDB中电影有链接关系的搜索。检测搜索是否指向几个著名的游戏网站 | 改进的多元线性回归模型 | 搜索数据对销量变化有指示意义，但对现有模型的预测性能提升较小，因此更适用于其它数据来源缺失的情况 |
| Asur, S., & Huberman, B. (2010) | 电影卖座率 | 通过Twitter Search Api获取3个月的与某些电影相关的全部微博，关键字为电影标题中所有词汇 | 改进的多元线性回归模型 | 拟合优度R2达到0.9以上，社交媒体的讨论情况对于电影销量有很好的指示作用 |
| Chevalier, J., & Mayzlin, D. (2006) | 亚马逊网站的和BN.com售书情况 | 亚马逊和BN.com的图书销售排名，图书评价，和上架时间、价格等基本信息 | 改进的多元线性回归模型 | 购书用户评价与图书销量排名之间的关系具有一定的相关性，在不同平台上关系不同 |
| Choi H, Varian H（2012） | 汽车销售、房屋销售 | 美国国家统计局公布的机动车及其零部件销售情况，Google 搜索与机动车有关的数据 | 自回归模型（Autoregressive models） | 加入搜索数据后模型在实时预报上有性能提升，样本外的相对提升达到21.5% |
| Barreira, N., Godinho, P., & Melo, P. (2013) | 实时监测预测汽车销量 | 四个国家的与汽车销量相关的Google 搜索数据 | 基本的自回归模型（Autoregressive models） | 部分情况下搜索数据可以帮助解释销量的方差，在实时预报上有优势 |
|  |  |  |  |  |
|  |  |  |  |  |
| Jacques Bughin | 实时监测预测比利时国家电信公司销量（网络和数字电视的） | Google 搜索数据与社交媒体数据的融合（Twitter, Facebook和其它博客数据） | 加入了短期预测变量的误差修正机制模型(Error Correction Mechanism，ECM) | 销量的 15%可以通过网络社交媒体数据解释，25%可以通过网络搜索数据解释；加入网络社交媒体和网络搜索数据以后，模型的整体预测能力提高了 25%。 |

文化产业包括电影、游戏和唱片等

销量预测主要使用的数据有销售大数据，搜索大数据，工业大数据，以及与宏观经济相关的数据等。方法为有较为经典的时间序列分析、灰色模型、也有较新的方法如机器学习、深度学习等。

搜索大数据在销量预测中能够取得较好的作用，原因在于网络搜索数据记录了数以亿计的搜索关注与需求，反映了市场主体的行为趋势与规律，可以为研究社会经济行为提供数据基础。例如消费者在购买汽车过程中，首先会有需求准备期，在需求准备的这段时间内，消费者会根据自身的喜好、汽车的价格和性能、以及自身的经济状况来获取潜在购买相关汽车的信息。随着网络技术的快速发展和互联网的普及，消费者更加青睐与从互联网上搜索相关数据和信息，因此而留下的记录可供销量预测等研究使用。

2.4产业发展水平的监测预测