

大数据择时看多，关注商业贸易、非银金融等行业

互联网大数据周报(20170611)

报告摘要:

● 网络大数据之择时

本周沪深 300 累计上涨 2.57%，大数据择时策略本周取得正收益，大数据择时策略在沪深 300 上的最新择时观点为看多。

本周具体表现及历史表现请见正文。

● 行业表现与网络舆情

本周表现最好前五行业：家用电器、食品饮料、建筑材料、计算机、电子

本周表现最差前五行业：公用事业、房地产、采掘、钢铁、银行

本周舆情上升最快前五行业：商业贸易、非银金融、食品饮料、化工、电气设备

本周舆情下降最快前五行业：公用事业、休闲服务、家用电器、建筑装饰、房地产

本周具体行业表现与舆情表现请见正文。

● 概念板块表现与网络舆情

本周表现最好前五概念板块：粤港澳自贸区、高送转概念、京津冀一体化、传感器、网络彩票

本周表现最差前五概念板块：民营医院、中日韩自贸区、移动互联网、油气改革、赛马

本周舆情上升最快前五概念板块：粤港澳自贸区、移动支付、国企改革、长江经济带、一带一路

本周舆情下降最快前五概念板块：海绵城市、油气改革、智能电视、迪士尼、创投

本周具体概念指数表现与舆情表现请见正文。

● 核心假设风险

周报中的策略经过历史数据建模、测算完成，网络数据不断更新，市场受众多因素影响，在复杂系统下存在失效可能性风险。

图 1 沪深300舆情走势

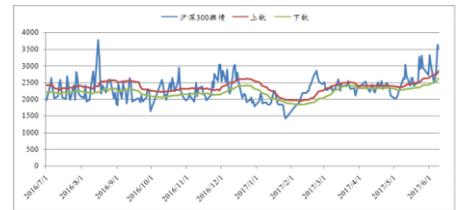


图 2 大数据舆情沪深300择时



分析师： 史庆盛 S0260513070004



02087577060



sqs@gf.com.cn

相关研究:

《多维数据下的大数据择时策略研究互联网大数据挖掘系列专题(八)》 2016-05-17

基于大数据挖掘的概念轮动策略-互联网大数据挖掘系列专题之(九) 2016-08-22

联系人： 陈原文 020-87576976

chenyuanwen@gf.com.cn

目录索引

一、 广发金工大数据挖掘框架简述	5
二、 网络大数据之择时	6
本周表现	7
历史表现	7
三、 行业表现与网络情绪	8
四、 概念板块表现与舆情	14
五、 大数据基金表现	16
六、 风险提示	17

图表索引

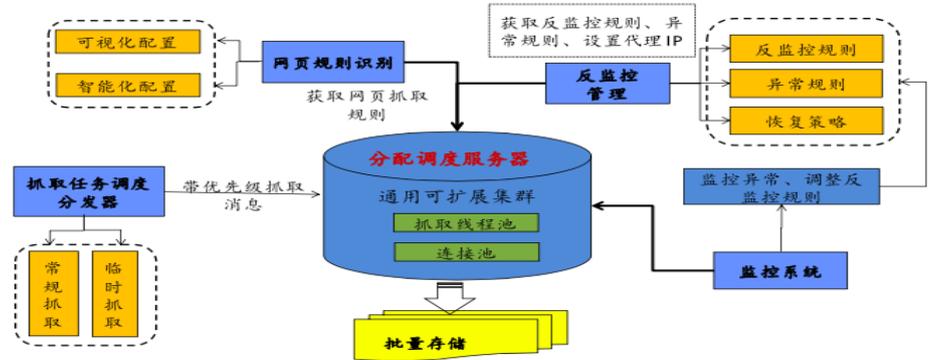
图 1 沪深 300 舆情走势.....	1
图 2 大数据舆情沪深 300 择时.....	1
图 3 广发金工数据抓取体系.....	5
图 4 广发金工大数据挖掘策略框架.....	5
图 5 沪深 300 舆情走势.....	7
图 6 沪深 300 舆情择时净值表现一览.....	8
图 7 申万一级行业过去一周表现.....	9
图 8 申万一级行业过去一周舆情表现.....	9
图 9 采掘(申万)行业舆情走势一览.....	9
图 10 传媒(申万)行业舆情走势一览.....	9
图 11 电气设备(申万)行业舆情走势一览.....	10
图 12 电子(申万)行业舆情走势一览.....	10
图 13 房地产(申万)行业舆情走势一览.....	10
图 14 纺织服装(申万)行业舆情走势一览.....	10
图 15 非银金融(申万)行业舆情走势一览.....	10
图 16 钢铁(申万)行业舆情走势一览.....	10
图 17 公用事业(申万)行业舆情走势一览.....	11
图 18 国防军工(申万)行业舆情走势一览.....	11
图 19 化工(申万)行业舆情走势一览.....	11
图 20 机械设备(申万)行业舆情走势一览.....	11
图 21 计算机(申万)行业舆情走势一览.....	11
图 22 家用电器(申万)行业舆情走势一览.....	11
图 23 建筑材料(申万)行业舆情走势一览.....	12
图 24 建筑装饰(申万)行业舆情走势一览.....	12
图 25 交通运输(申万)行业舆情走势一览.....	12
图 26 农林牧渔(申万)行业舆情走势一览.....	12
图 27 汽车(申万)行业舆情走势一览.....	12
图 28 轻工制造(申万)行业舆情走势一览.....	12
图 29 商业贸易(申万)行业舆情走势一览.....	13
图 30 食品饮料(申万)行业舆情走势一览.....	13
图 31 通信(申万)行业舆情走势一览.....	13
图 32 休闲服务(申万)行业舆情走势一览.....	13
图 33 医药生物(申万)行业舆情走势一览.....	13
图 34 银行(申万)行业舆情走势一览.....	13
图 35 有色金属(申万)行业舆情走势一览.....	14
图 36 综合(申万)行业舆情走势一览.....	14
图 37 粤港澳自贸区舆情走势一览.....	15
图 38 高送转舆情走势一览.....	15
图 39 京津冀一体化舆情走势一览.....	15

图 40 传感器舆情走势一览.....	15
图 41 网络彩票舆情走势一览.....	15
图 42 赛马舆情走势一览.....	15
图 43 油气改革舆情走势一览.....	16
图 44 移动互联网舆情走势一览.....	16
图 45 中日韩自贸区舆情走势一览.....	16
图 46 民营医院舆情走势一览.....	16
表 1 大数据相关策略研究成果.....	6
表 2 互联网挖掘小工具一览.....	6
表 3 大数据择时本周表现(2017.06.05 -2017.06.09).....	7
表 4 大数据择时历史表现一览.....	7
表 5 过去一周行业表现及舆情表现一览.....	8
表 6 过去一周概念板块表现及舆情表现一览.....	14
表 7 大数据基金表现一览.....	16

一、 广发金工大数据挖掘框架简述

在互联网数据挖掘与研究方面，广发金工建立了完善的数据抓取平台以及对互联网大数据各个数据维度的研究视角。

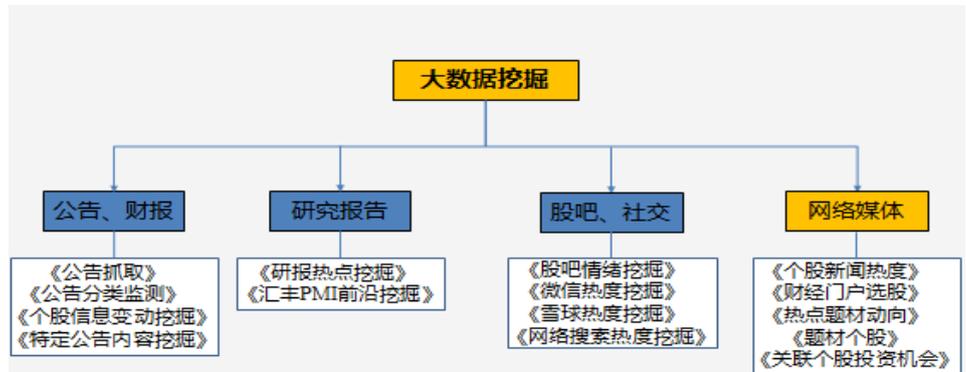
图 3 广发金工数据抓取体系



数据来源：广发证券发展研究中心

广发金工在大数据挖掘体系下，对整个市场的多个维度的数据进行了深入挖掘，包括卖方研究报告、上市公司公告、股吧舆情、个股关注度以及热度新闻等多个领域。

图 4 广发金工大数据挖掘策略框架



数据来源：广发证券发展研究中心

广发金工研究了多个维度的互联网数据，并对相关维度的数据进行了深度的研究，构建了相关的策略。比如从公告、财报角度研究的报告《公告披露背后隐藏的投资机会——互联网大数据挖掘系列专题之（二）》等；从股吧、社交角度研究的报告《倾听股吧之声，洞察大盘趋势——互联网大数据挖掘系列专题之（三）》等；以及从网络媒体角度分析的报告《基于互联网挖掘的热点选股策略——互联网大数据挖掘系列专题之（五）》、《基于大数据挖掘的关联个股投资机会——互联网大数据挖掘系列专题之（六）》、《基于大数据挖掘的Smart Beta策略——互联网大数据挖掘系列专题之（七）》、《多维数据下的大数据择时策略研究——互联网大数据挖掘系列专题之（八）》、《基于大数据挖掘的概念轮动策略-互联网大数据挖掘系列专题之（九）》等。

表 1 大数据相关策略研究成果

《基于网络新闻热度的择时策略——互联网大数据挖掘系列专题(一)》
《公告披露背后隐藏的投资机会——互联网大数据挖掘系列专题之(二)》
《倾听股吧之声，洞察大盘趋势——互联网大数据挖掘系列专题之(三)》
《那些年一起追过的财经小编选股策略——互联网大数据挖掘系列专题之(四)》
《基于互联网挖掘的热点选股策略——互联网大数据挖掘系列专题之(五)》
《基于大数据挖掘的关联个股投资机会——互联网大数据挖掘系列专题之(六)》
《基于大数据挖掘的 Smart Beta 策略——互联网大数据挖掘系列专题之(七)》
《多维数据下的大数据择时策略研究——互联网大数据挖掘系列专题之(八)》
《基于大数据挖掘的概念轮动策略-互联网大数据挖掘系列专题之(九)》

数据来源：广发证券发展研究中心

基于对大数据的研究积累，广发金工也开发了一系列的互联网挖掘小工具，包括上市公司公告抓取、研究报告抓取、搜索量抓取、关注度抓取等小工具。

表 2 互联网挖掘小工具一览

A 股新闻热度搜索工具	A 股上市工具公告抓取工具
上市公司信息变更抓取	文本信息批量识别及处理
汇丰 PMI 实时监测工具	个股研报热点监测工具
特定公告实时监测工具	财经小编选股工具

数据来源：广发证券发展研究中心

本《互联网大数据周报》将利用广发金工积累的丰富的网络数据，并结合A股市场的行情数据等，在广发金工原有的研究成果的基础上，对舆情数据在择时、选股、行业配置、概念轮动等领域进行跟踪，给投资者对当前的市场的走势、选股、行业配置、概念轮动等领域提供参考。

二、网络大数据之择时

股市的涨跌通常伴随着成交额的增减变动，而散户投资者作为A股成交额的主力军，他们的情绪对市场涨跌有着推波助澜的作用，是一种领先于市场行情的指标。当投资者对市场情绪较乐观时，会先搜索、关注相关标的的财务数据、行业信息等相关信息，然后进行投资交易，而行情数据是投资者交易的结果，已经反映了这些情绪的变动，是相对于投资者情绪的滞后数据。

我们基于一些热门专业的财经网站和搜索引擎，比如百度搜索引擎、东方财富网、新浪财经、雪球网、和讯财经等互联网数据源，利用网站上股票的关注人数、搜索量数据，得到关于大盘的情绪指标，利用这些舆情指标探讨网络热度数据在大盘指数上的择时应用。

基于广发金工团队对大数据舆情在择时研究中的积累，得到的研究结论是：当当日网络舆情指标较大时，次日股市上涨的可能性较大，当当日网络舆情指标较小时，次日股市下跌的可能性较大。只有当网络舆情在短期内积累到一定程度时，次日股市上涨才会成为一个可预测的事件，同理，只有当网络舆情在短期内下跌到一定程度时，次日股市的下跌的可能性才更大。为了更好地描述这一过程，对网络舆情指标设置一个阈值，高于或者低于某个阈值，才发出看多或者看空的信号。为了更够动态地描述网络舆情数据的变化，因此，需要寻找到不同时点上的阈值，才能

准确地运用网络舆情数据预测次日市场的涨跌走势。

策略原理:

设置布林通道: 以热度数据的M日均值构造均线, 以M日热度数据的标准差的N倍构造上轨、下轨, 即:

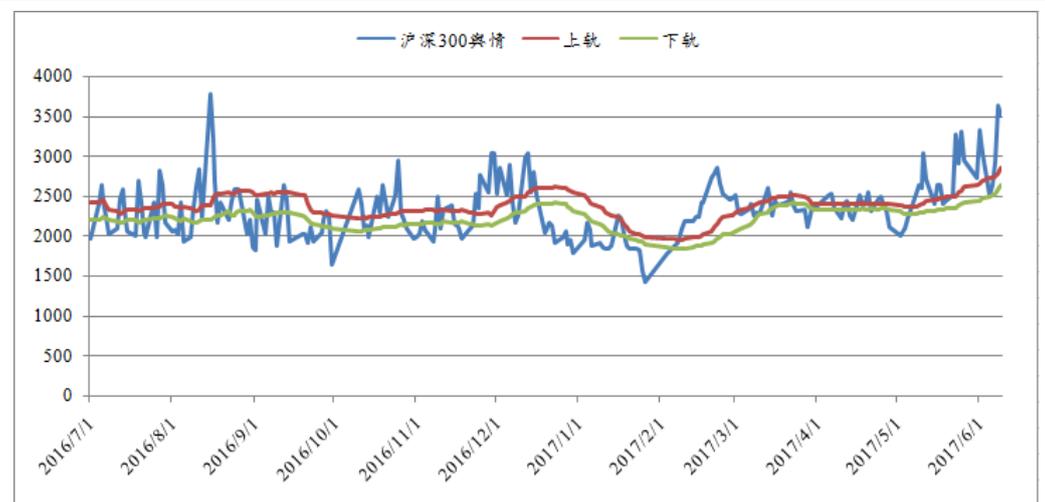
布林通道上轨=热度数据的M日均值+N*M日热度的标准差;

布林通道下轨=热度数据的M日均值-N*M日热度的标准差;

择时策略: 当日热度突破通道上轨, 则发出看多信号, 次日买入; 当日热度突破通道下轨, 则发出看空信号, 次日卖出; 相连两日发出不同信号, 先平仓后根据信号进行买卖; 相连两日的第二日不发出信号, 次日开盘直接平仓, 考虑可以做空的收益;

注: 本周报跟踪的是利用网络舆情数据对沪深300指数的择时跟踪。

图 5 沪深300舆情走势



数据来源: Wind、广发证券发展研究中心

本周表现

表 3 大数据择时本周表现(2017.06.05 -2017.06.09)

本周收益率	0.68%	看多次数	3
本周盈亏	盈	看多胜率	100.00%
最大回撤	0.00%	看空次数	0
胜率	40.00%	看空胜率	#DIV/0!

数据来源: 广发证券发展研究中心

历史表现

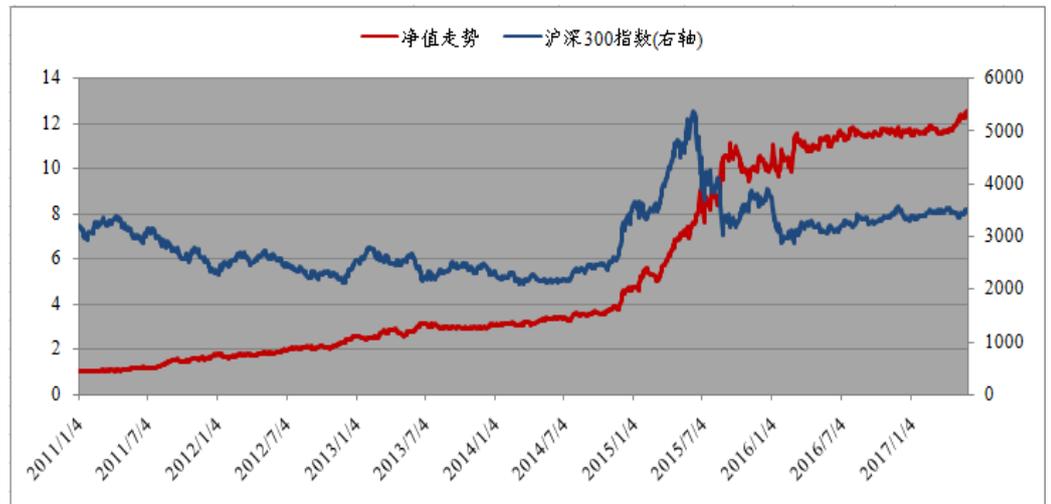
表 4 大数据择时历史表现一览

分年度指标统计	累计收益率	最大回撤	胜率	看多次数	看多胜率	看空次数	看空胜率	年度盈亏
2011	71.82%	-6.41%	61.49%	81	55.56%	92	67.39%	盈
2012	47.78%	-10.31%	56.46%	97	57.73%	112	55.36%	盈
2013	18.11%	-10.73%	55.38%	70	58.57%	116	53.45%	盈

2014	55.79%	-4.72%	58.76%	110	60.00%	84	57.14%	盈
2015	115.53%	-18.11%	55.50%	92	61.96%	117	50.43%	盈
2016	15.65%	-12.47%	51.26%	65	55.38%	134	49.25%	盈
2017 年至今	7.63%	-3.37%	59.76%	46	63.04%	36	55.56%	盈
整体	1153.31%	-18.11%	56.63%	561	58.82%	691	54.85%	盈

数据来源：广发证券发展研究中心

图 6 沪深300与情择时净值表现一览



数据来源：Wind、广发证券发展研究中心

三、 行业表现与网络情绪

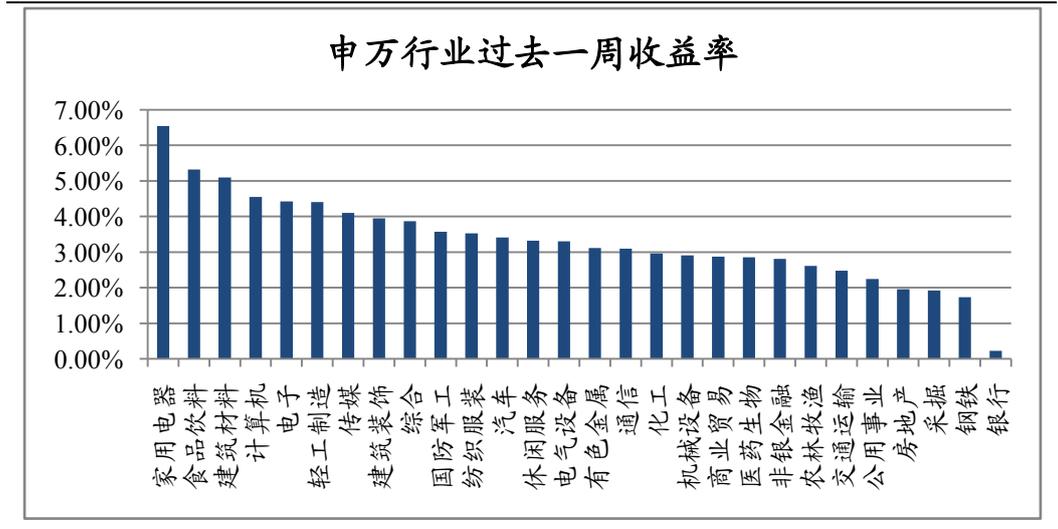
本部分内容主要是根据申万一级行业分类，跟踪过去一周申万一级行业指数走势与网络舆情之间的关系。

表 5 过去一周行业表现及舆情表现一览

本周表现最好前五行业	本周表现最差前五行业
家用电器:6.54%	公用事业:2.24%
食品饮料:5.32%	房地产:1.96%
建筑材料:5.10%	采掘:1.92%
计算机:4.55%	钢铁:1.73%
电子:4.42%	银行:0.23%
本周舆情上升最快前五行业	本周舆情下降最快前五行业
商业贸易:90.91%	公用事业:-11.06%
非银金融:9.16%	休闲服务:-11.48%
食品饮料:7.91%	家用电器:-11.94%
化工:5.92%	建筑装饰:-20.00%
电气设备:4.97%	房地产:-26.20%

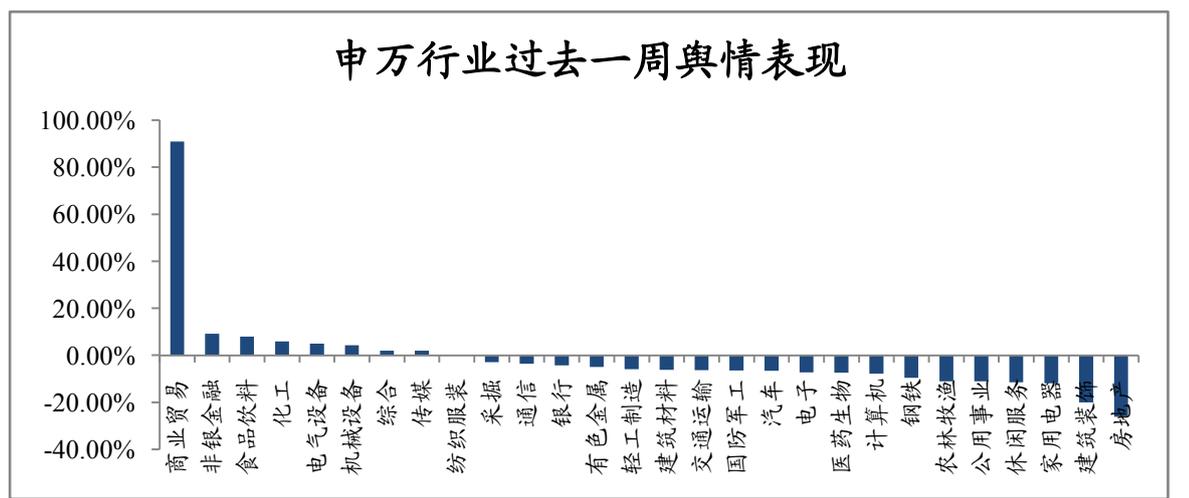
数据来源：Wind、广发证券发展研究中心

图 7 申万一级行业过去一周表现



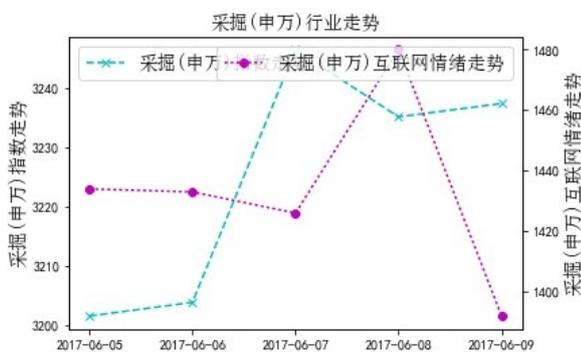
数据来源: Wind、广发证券发展研究中心

图 8 申万一级行业过去一周舆情表现



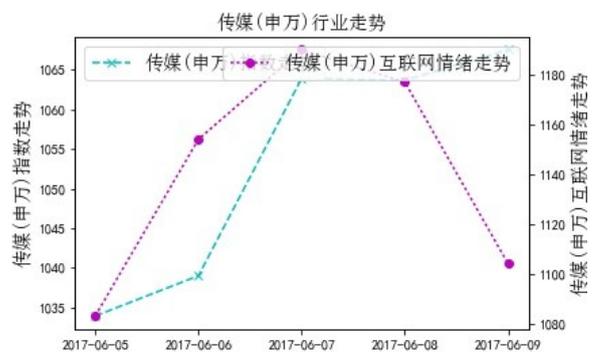
数据来源: Wind、广发证券发展研究中心

图 9 采掘(申万)行业舆情走势一览



数据来源: Wind、广发证券发展研究中心

图 10 传媒(申万)行业舆情走势一览



数据来源: Wind、广发证券发展研究中心

图 11 电气设备(申万)行业舆情走势一览

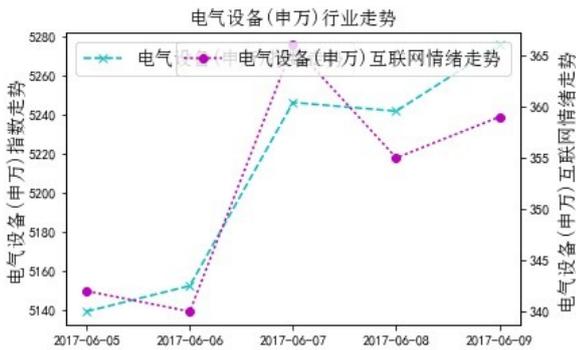
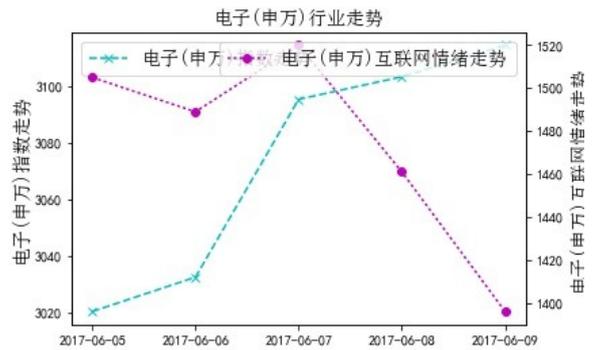


图 12 电子(申万)行业舆情走势一览



数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 13 房地产(申万)行业舆情走势一览

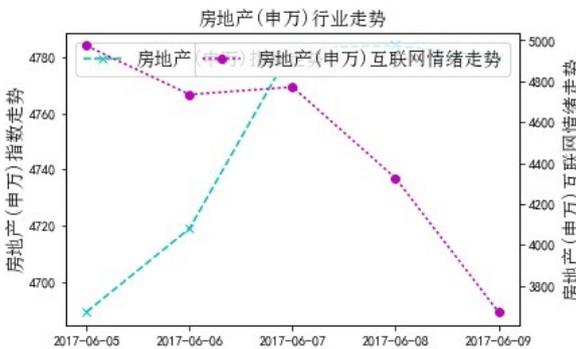
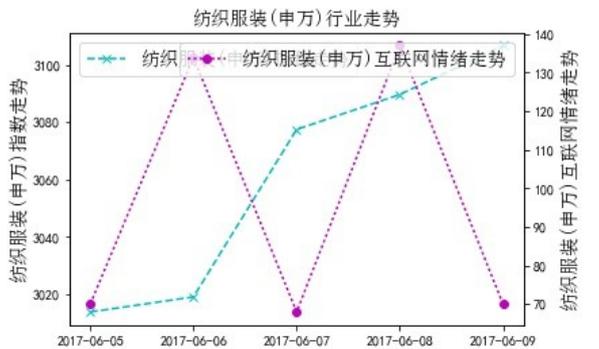


图 14 纺织服装(申万)行业舆情走势一览



数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 15 非银金融(申万)行业舆情走势一览

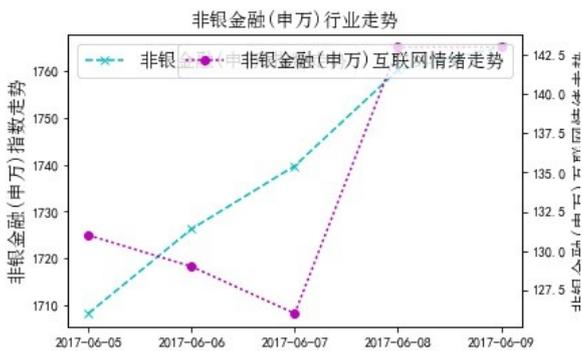
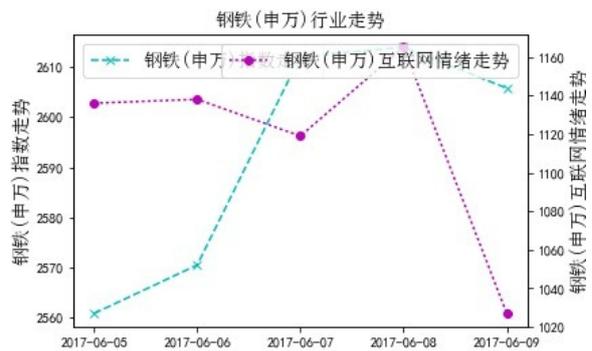


图 16 钢铁(申万)行业舆情走势一览

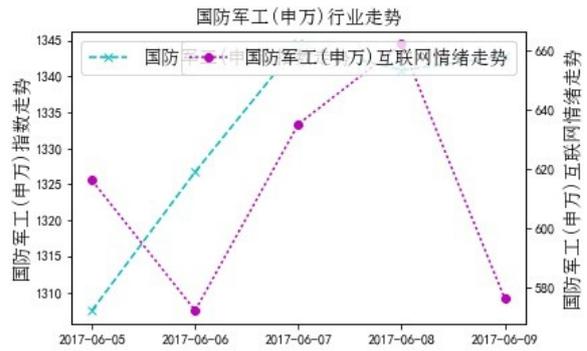
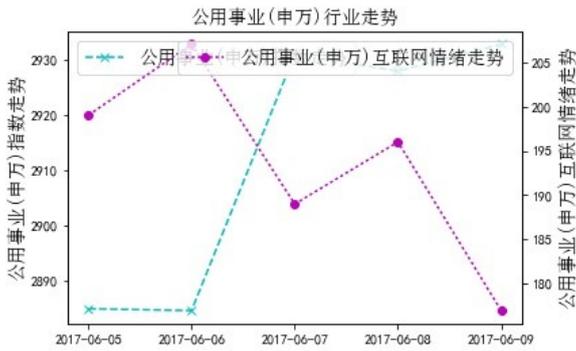


数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 17 公用事业(申万)行业舆情走势一览

图 18 国防军工(申万)行业舆情走势一览

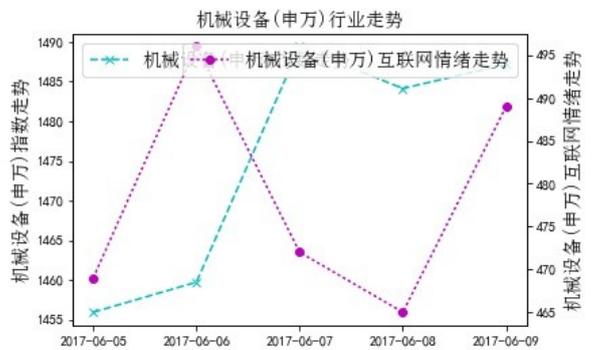
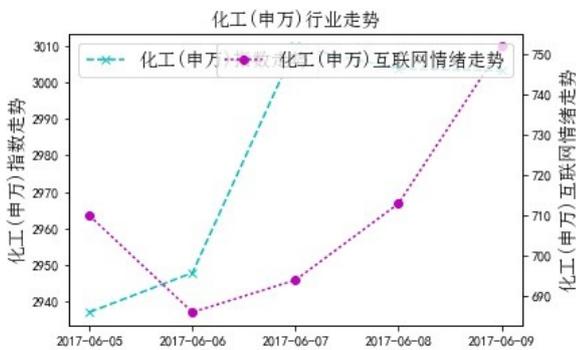


数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 19 化工(申万)行业舆情走势一览

图 20 机械设备(申万)行业舆情走势一览

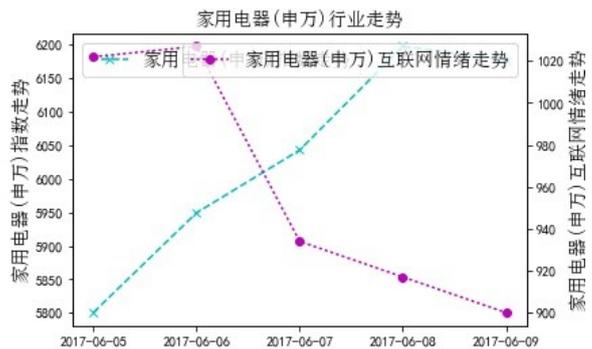
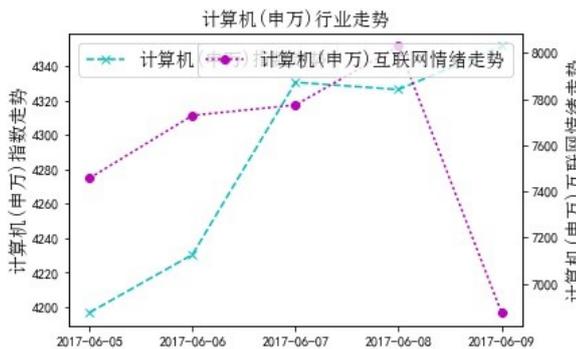


数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 21 计算机(申万)行业舆情走势一览

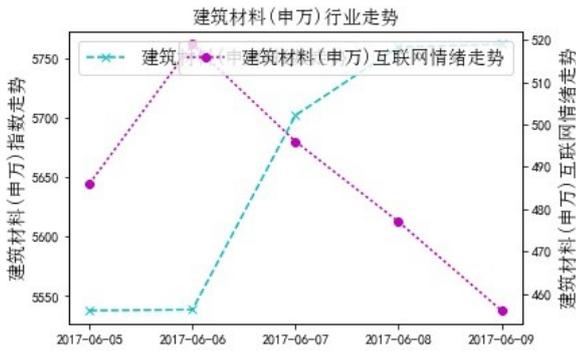
图 22 家用电器(申万)行业舆情走势一览



数据来源: Wind、广发证券发展研究中心

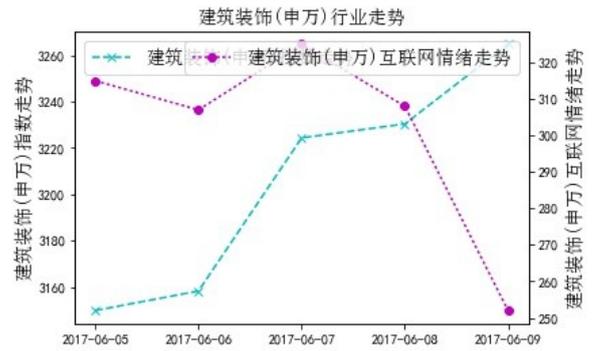
数据来源: Wind、广发证券发展研究中心

图 23 建筑材料(申万)行业舆情走势一览



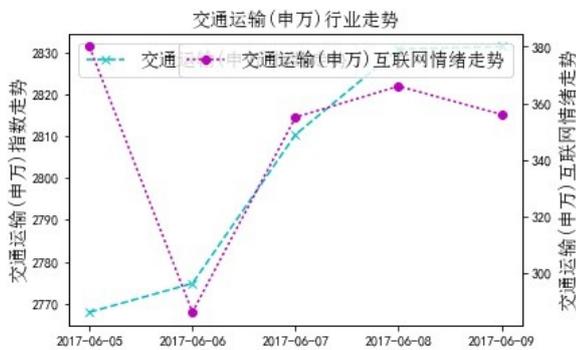
数据来源: Wind、广发证券发展研究中心

图 24 建筑装饰(申万)行业舆情走势一览



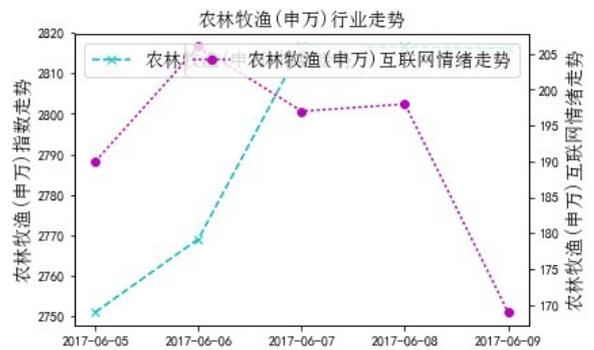
数据来源: Wind、广发证券发展研究中心

图 25 交通运输(申万)行业舆情走势一览



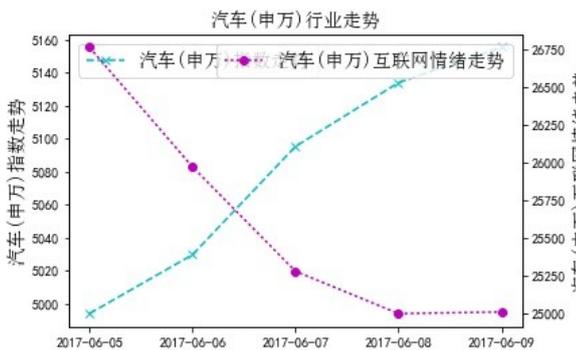
数据来源: Wind、广发证券发展研究中心

图 26 农林牧渔(申万)行业舆情走势一览



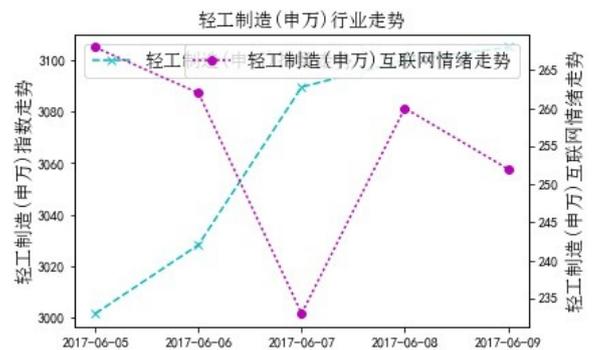
数据来源: Wind、广发证券发展研究中心

图 27 汽车(申万)行业舆情走势一览



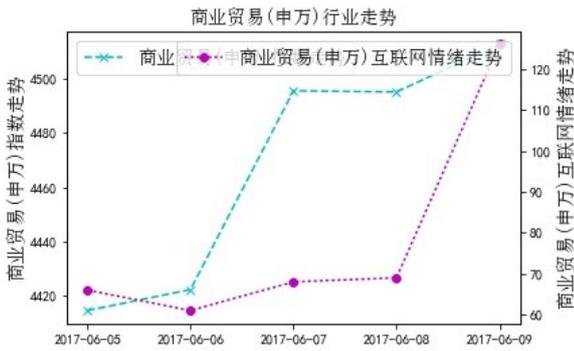
数据来源: Wind、广发证券发展研究中心

图 28 轻工制造(申万)行业舆情走势一览



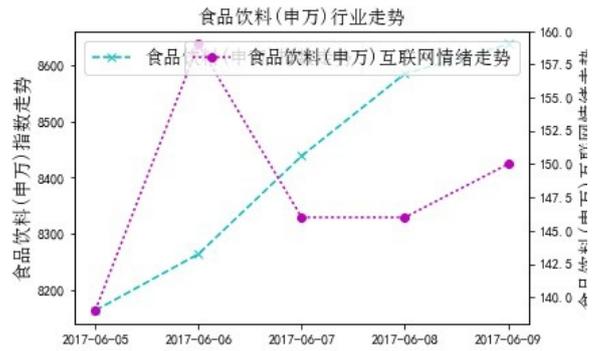
数据来源: Wind、广发证券发展研究中心

图 29 商业贸易(申万)行业舆情走势一览



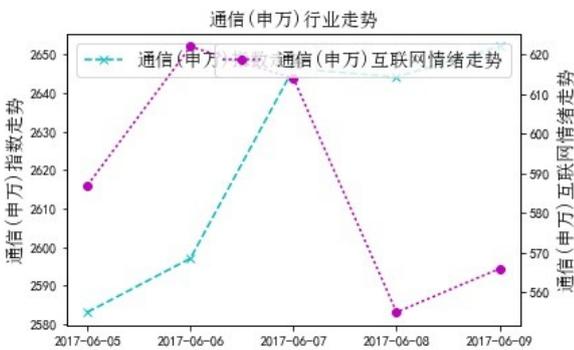
数据来源: Wind、广发证券发展研究中心

图 30 食品饮料(申万)行业舆情走势一览



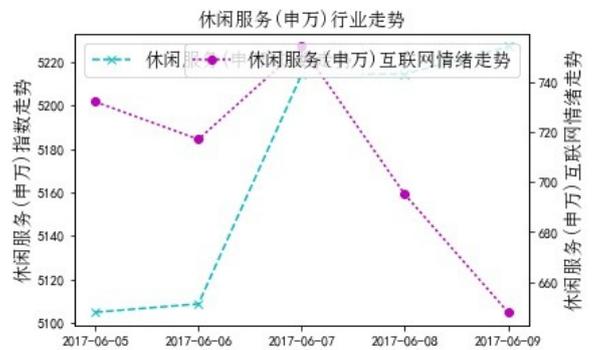
数据来源: Wind、广发证券发展研究中心

图 31 通信(申万)行业舆情走势一览



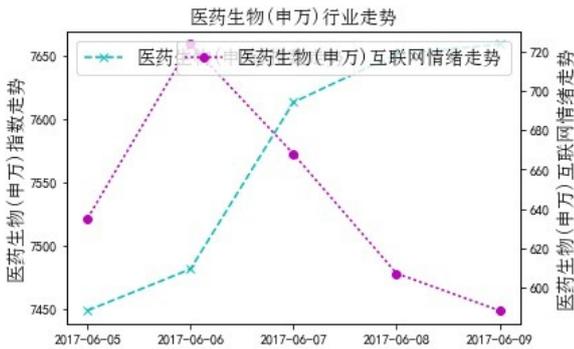
数据来源: Wind、广发证券发展研究中心

图 32 休闲服务(申万)行业舆情走势一览



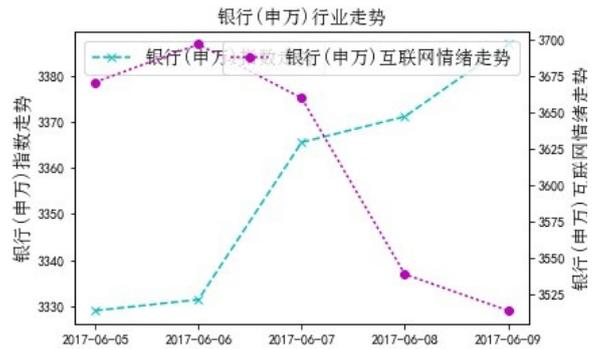
数据来源: Wind、广发证券发展研究中心

图 33 医药生物(申万)行业舆情走势一览



数据来源: Wind、广发证券发展研究中心

图 34 银行(申万)行业舆情走势一览



数据来源: Wind、广发证券发展研究中心

图 35 有色金属(申万)行业舆情走势一览

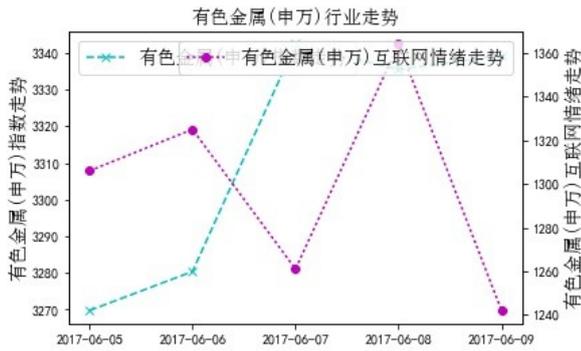
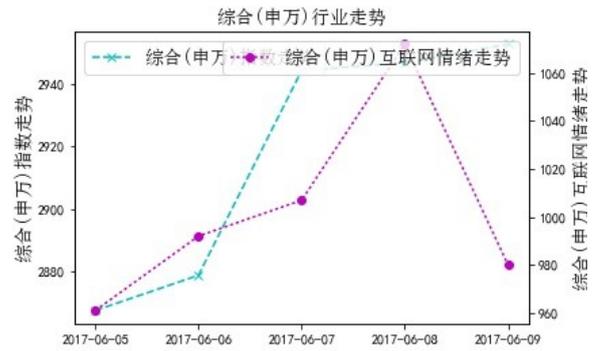


图 36 综合(申万)行业舆情走势一览



数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

四、概念板块表现与舆情

近年来, A股市场上概念、热点的炒作频繁, 例如今年来的“深港通”、“国企改革”、“互联网金融”、“OLED”等, 相关个股在某一概念、热点背景下, 往往在中短期时间内有较好的市场表现。如何抓住这种热点、概念带来的投资机会, 是当前量化研究的一个难点。本节内容主要统计过去1周wind概念板块指数中表现最好以及最差的几个概念板块的行情表现以及对应的网络舆情表现。

表 6 过去一周概念板块表现及舆情表现一览

涨幅前五概念指数	本周行情表现	本周舆情表现
粤港澳自贸区	9.16%	113.24%
高送转概念	6.78%	4.24%
京津冀一体化	5.74%	-3.79%
传感器	5.69%	-2.84%
网络彩票	5.53%	-8.00%
跌幅前五概念指数	本周行情表现	本周舆情表现
赛马	1.47%	-12.48%
油气改革	1.90%	-21.30%
移动互联网	1.99%	1.52%
中日韩自贸区	2.07%	-3.23%
民营医院	2.13%	0.62%

数据来源: Wind、广发证券发展研究中心

图 37 粤港澳自贸区行情走势一览

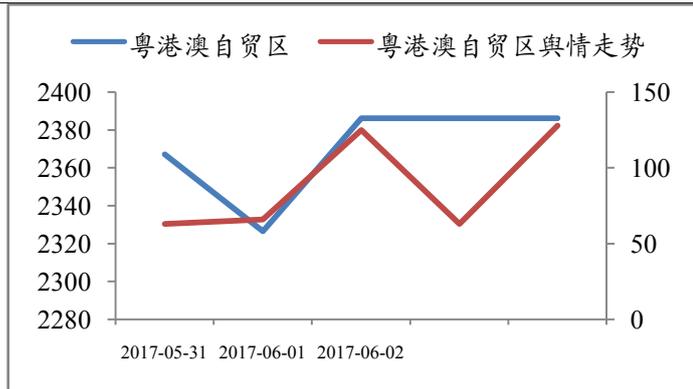
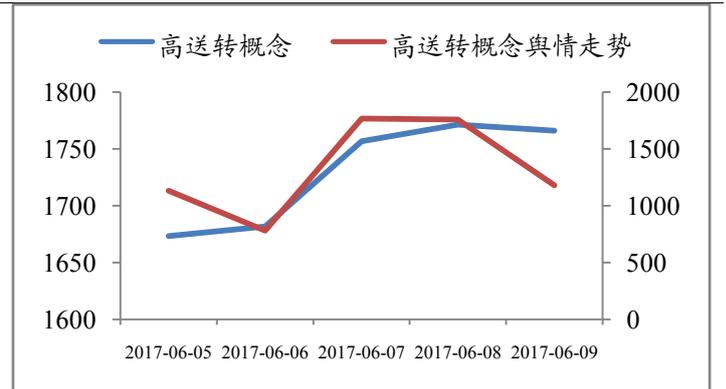


图 38 高送转行情走势一览



数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 39 京津冀一体化行情走势一览

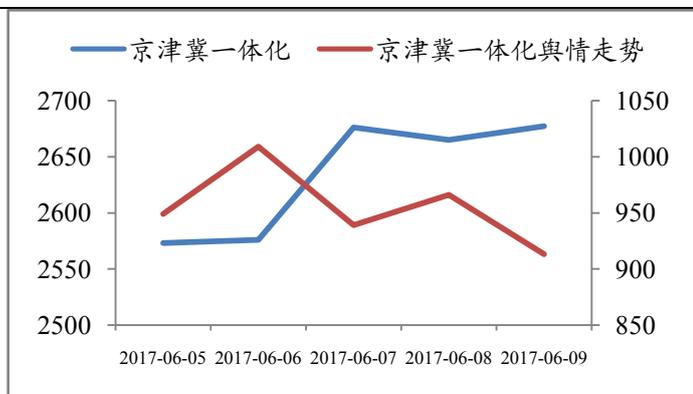
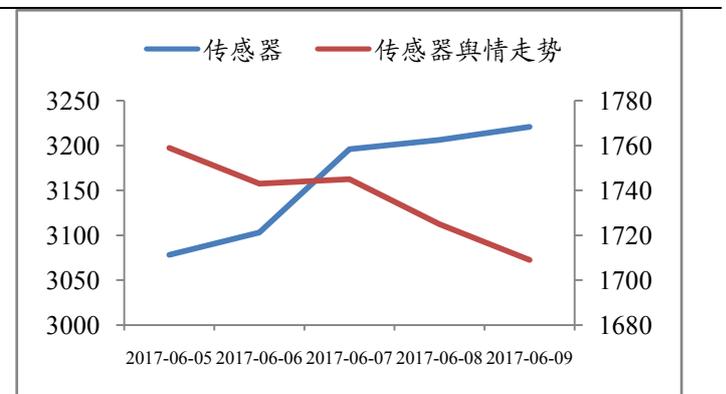


图 40 传感器行情走势一览



数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 41 网络彩票行情走势一览

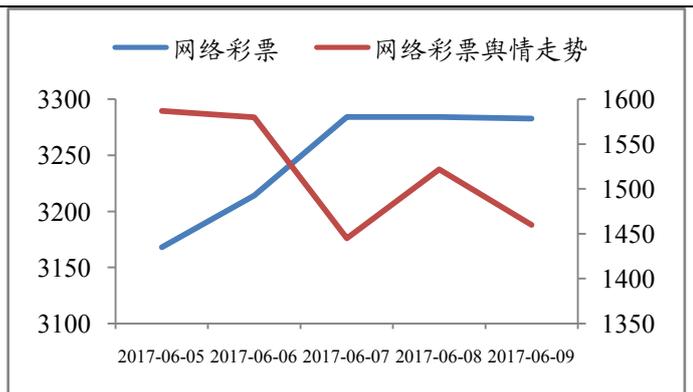
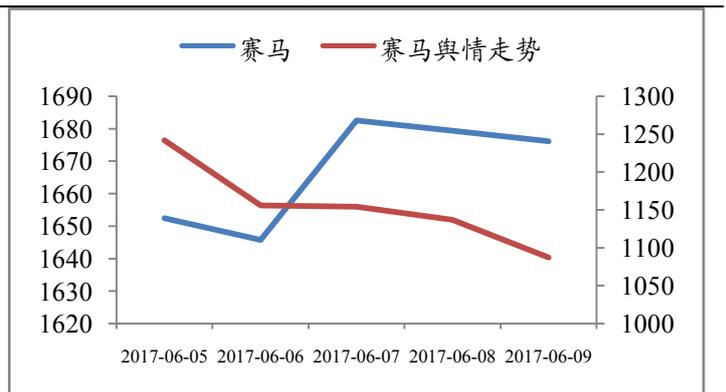


图 42 赛马行情走势一览



数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 43 油气改革行情走势一览

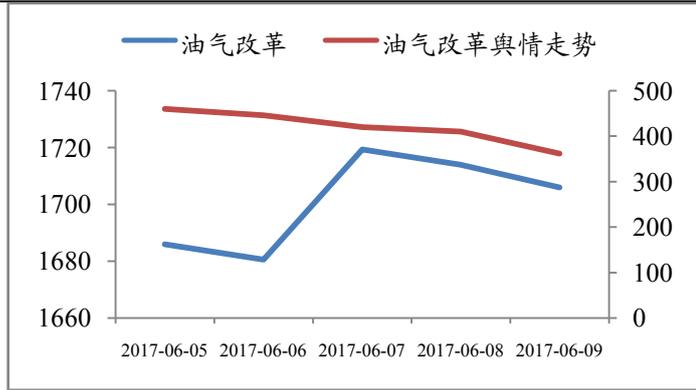


图 44 移动互联网行情走势一览



数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

图 45 中日韩自贸区行情走势一览

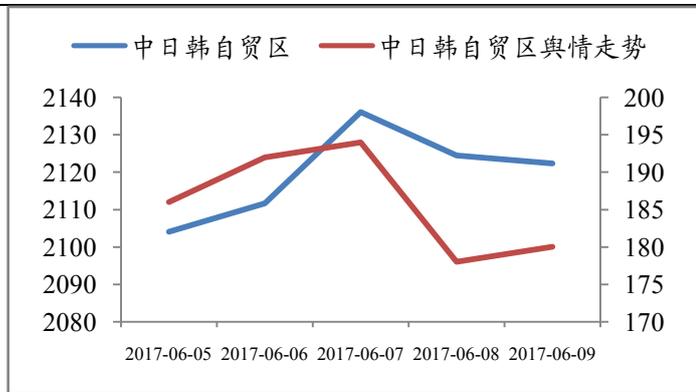
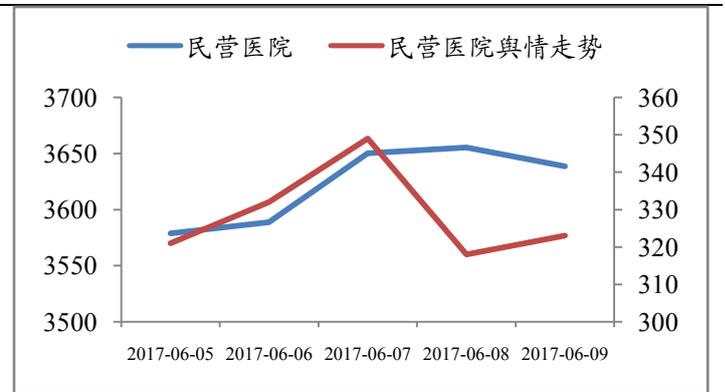


图 46 民营医院行情走势一览



数据来源: Wind、广发证券发展研究中心

数据来源: Wind、广发证券发展研究中心

五、 大数据基金表现

表 7 大数据基金表现一览

基金代码	基金名称	基金成立以来表现	最近一周表现	最近一个月表现	最大回撤	夏普比率
002465.OF	东兴众智优选	-3.10%	2.98%	3.97%	-12.87%	-10.09
001242.OF	博时淘金大数据 100A	-14.15%	2.90%	4.30%	-53.98%	13.52
001243.OF	博时淘金大数据 100I	-14.13%	2.90%	4.30%	-53.97%	13.53
000827.OF	广发中证百发 100E	15.21%	2.76%	5.29%	-60.22%	-6.85
000826.OF	广发中证百发 100A	15.45%	2.75%	5.28%	-60.21%	-6.75
001426.OF	南方大数据 300C	6.44%	2.74%	3.19%	-24.72%	0.61
001420.OF	南方大数据 300A	7.04%	2.74%	3.19%	-24.69%	0.93
002588.OF	博时银智大数据 100	21.09%	2.73%	2.58%	-9.18%	3.46
001564.OF	东方红京东大数据	28.20%	2.72%	3.55%	-18.03%	12.14
001742.OF	广发百发大数据 E	-0.90%	2.59%	4.87%	-16.86%	-0.32
001741.OF	广发百发大数据 A	-0.90%	2.59%	4.87%	-16.94%	-0.32

001637.OF	嘉实腾讯自选股大数据	0.80%	2.44%	4.24%	-14.73%	-11.56
001030.OF	天弘云端生活优选	-25.20%	2.38%	3.67%	-59.00%	-4.75
519677.OF	银河定投宝	69.40%	2.29%	4.70%	-50.42%	2.24
002236.OF	大成互联网+大数据 A	10.20%	2.23%	5.25%	-18.58%	-0.26
519132.OF	海富通东财大数据	-5.00%	2.15%	2.70%	-17.16%	-16.01
002263.OF	泰达宏利同顺大数据 A	12.29%	1.87%	3.18%	-11.85%	2.88
003416.OF	招商财经大数据策略	-19.07%	1.81%	4.55%	-24.95%	-45.01
002967.OF	浙商大数据智选消费	-1.40%	1.75%	2.92%	-6.17%	-13.06
001113.OF	南方大数据 100A	-19.02%	1.73%	3.20%	-49.78%	11.95
002269.OF	银华大数据	-10.20%	1.13%	1.13%	-11.58%	-28.56
001734.OF	广发百发大数据策略成长 A	-7.10%	0.98%	2.20%	-14.97%	-2.61
001735.OF	广发百发大数据策略成长 E	-7.00%	0.98%	2.20%	-14.87%	-2.55
003055.OF	博时招财二号大数据	0.50%	0.60%	0.50%	-2.09%	2.98
001238.OF	博时招财一号大数据	2.90%	0.00%	0.00%	-3.48%	1.74

数据来源：Wind、广发证券发展研究中心

六、 风险提示

周报中的策略经过历史数据建模、测算完成，网络数据不断更新，市场受众多因素影响，在复杂系统下存在失效可能性风险。

广发证券—行业投资评级说明

- 买入： 预期未来 12 个月内，股价表现强于大盘 10%以上。
持有： 预期未来 12 个月内，股价相对大盘的变动幅度介于-10%~+10%。
卖出： 预期未来 12 个月内，股价表现弱于大盘 10%以上。

广发证券—公司投资评级说明

- 买入： 预期未来 12 个月内，股价表现强于大盘 15%以上。
谨慎增持： 预期未来 12 个月内，股价表现强于大盘 5%-15%。
持有： 预期未来 12 个月内，股价相对大盘的变动幅度介于-5%~+5%。
卖出： 预期未来 12 个月内，股价表现弱于大盘 5%以上。

联系我们

	广州市	深圳市	北京市	上海市
地址	广州市天河区林和西路9号耀中广场A座1401	深圳福田区益田路6001号太平金融大厦31楼	北京市西城区月坛北街2号月坛大厦18层	上海市浦东新区富城路99号震旦大厦18楼
邮政编码	510620	518000	100045	200120
客服邮箱	gfyf@gf.com.cn			
服务热线				

免责声明

广发证券股份有限公司具备证券投资咨询业务资格。本报告只发送给广发证券重点客户，不对外公开发布。

本报告所载资料的来源及观点的出处皆被广发证券股份有限公司认为可靠，但广发证券不对其准确性或完整性做出任何保证。报告内容仅供参考，报告中的信息或所表达观点不构成所涉证券买卖的出价或询价。广发证券不对因使用本报告的内容而引致的损失承担任何责任，除非法律法规有明确规定。客户不应以本报告取代其独立判断或仅根据本报告做出决策。

广发证券可发出其它与本报告所载信息不一致及有不同结论的报告。本报告反映研究人员的不同观点、见解及分析方法，并不代表广发证券或其附属机构的立场。报告所载资料、意见及推测仅反映研究人员于发出本报告当日的判断，可随时更改且不予通告。

本报告旨在发送给广发证券的特定客户及其它专业人士。未经广发证券事先书面许可，任何机构或个人不得以任何形式翻版、复制、刊登、转载和引用，否则由此造成的一切不良后果及法律责任由私自翻版、复制、刊登、转载和引用者承担。