

※本資料は平成29年7月時点のもの

SNS×AI 鉱工業生産予測指数のご紹介

平成28年度 IoTを活用した新ビジネス創出推進事業(ビッグデータを活用した新指標開発事業)

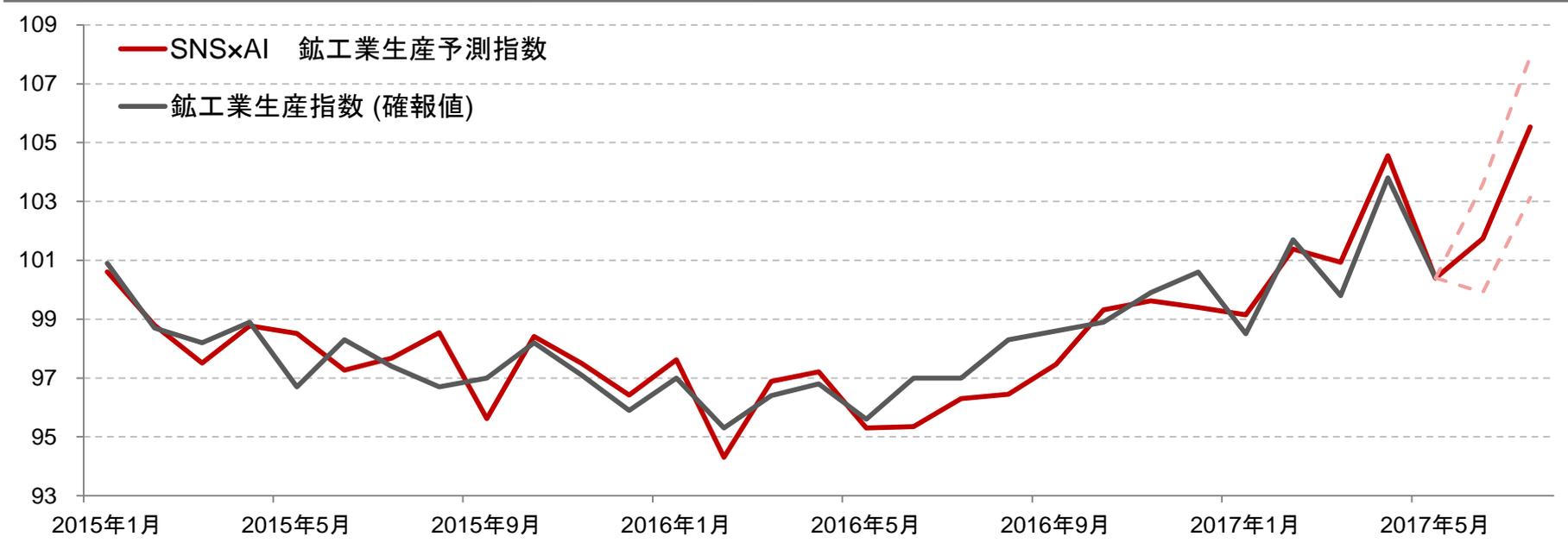
野村證券株式会社
金融工学研究センター

SNS×AI 鉱工業生産予測指数



SNS×AI 鉱工業生産予測指数

*各予測値は週次で上書き更新されます



*) 予測値の誤差は±2σで表示
出所)野村証券

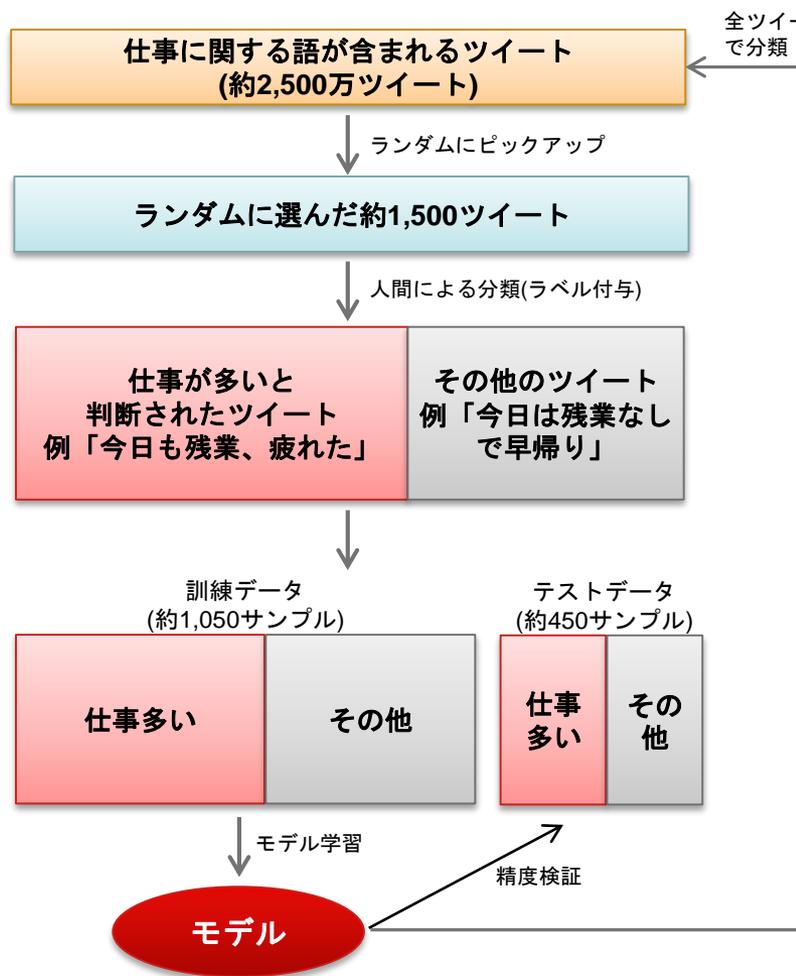
※本資料は平成29年7月時点のもの

NOMURA

SNS×AI 鉱工業生産予測指数：算出の仕組み

～ 仕事が増えたことを意味するツイートを抽出 ～

仕事に関するキーワードを含むツイートの中から、仕事が増えたことを意味するツイートのみを抽出するAIを開発
教師データとして、人間が約1,500のツイートを手動で分類し、それをAIに学習させています



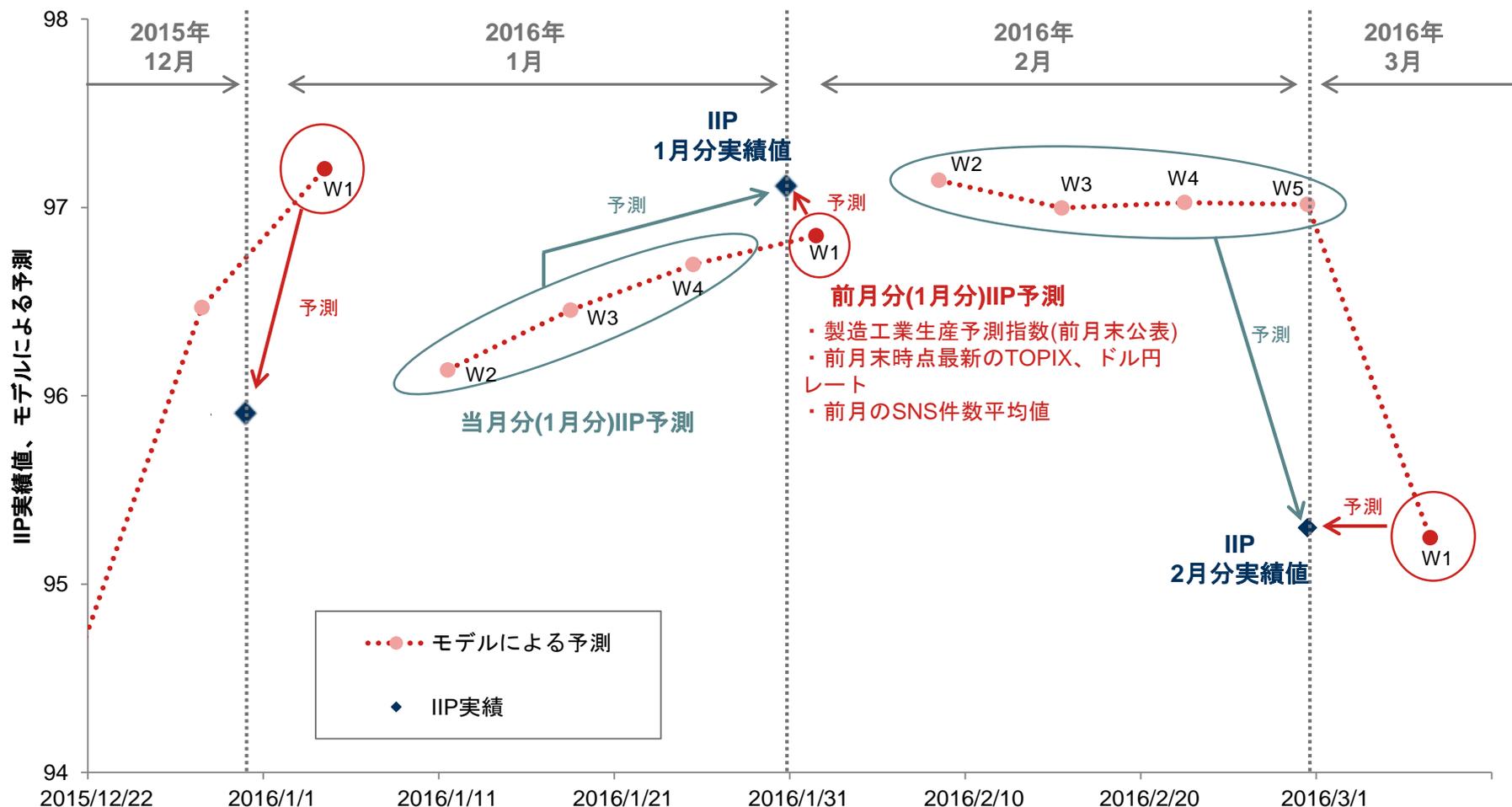
- 仕事に関するツイートから、仕事が増えたことを意味するツイートのみを抽出するAIを作成
 - ○: 今日も残業、疲れた
 - ×: 今日は残業無しで早帰り!
- テストデータの精度は約84.1%
 - 人間が付与する教師データを増やすことで、より精度を高められると考えられる

SNS × AI 鉱工業生産予測指数の週次予測方法

各月の第2週以降、当月の鉱工業生産指数(IIP)を予測します。

第1週はSNSデータが十分に取得できない可能性も考えられるため、前月分のIIPを予測します。

なお、予測にはSNSデータだけでなく、製造工業予測指数、為替レート、等のデータも併せて使用しています。



※本資料は平成29年7月時点のもの

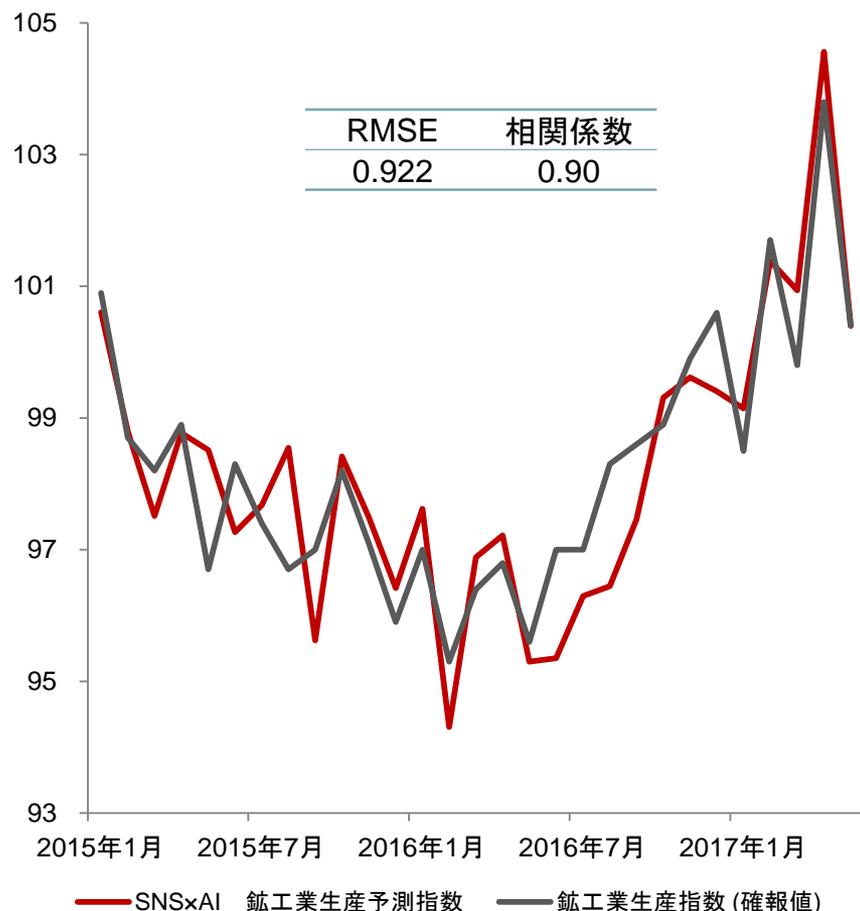
NOMURA

SNS×AI 鉱工業生産予測指数: 予測精度

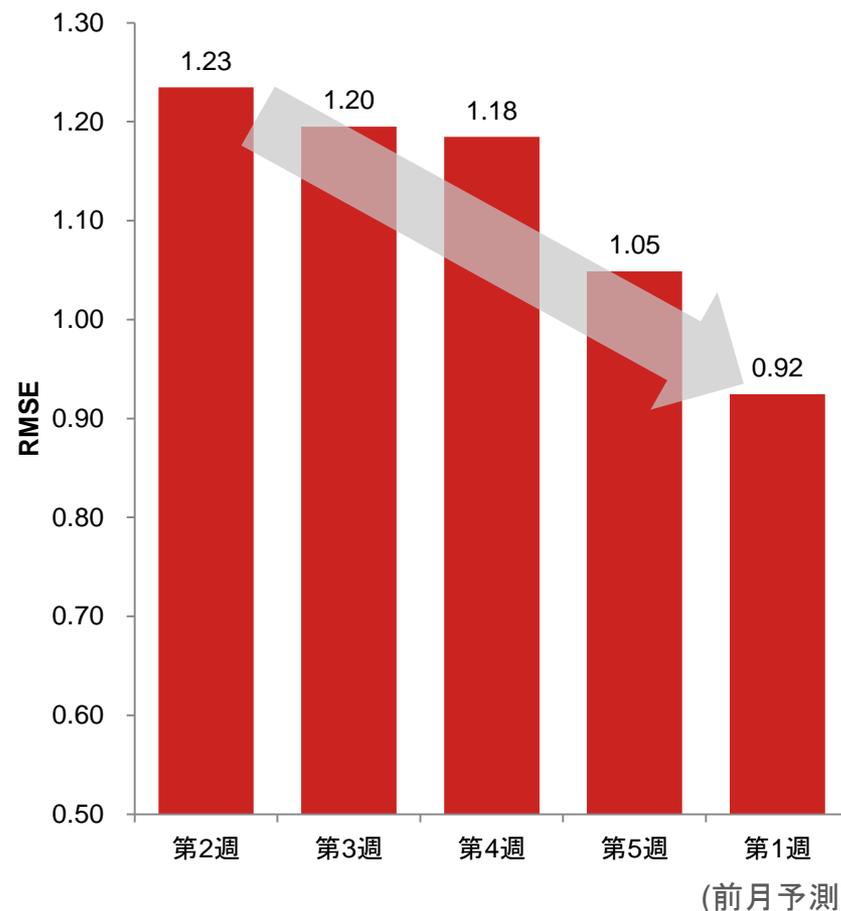
SNS × AI 鉱工業生産予測指数の週ごとの予測精度

第2週以降、週を重ねるごとにSNSデータが増加し、当月予測の精度は高まっています。
第1週での前月予測の精度が最も高くなっています。

前月予測とIIP実績値の時系列



予想期間別のRMSE(平均二乗誤差)

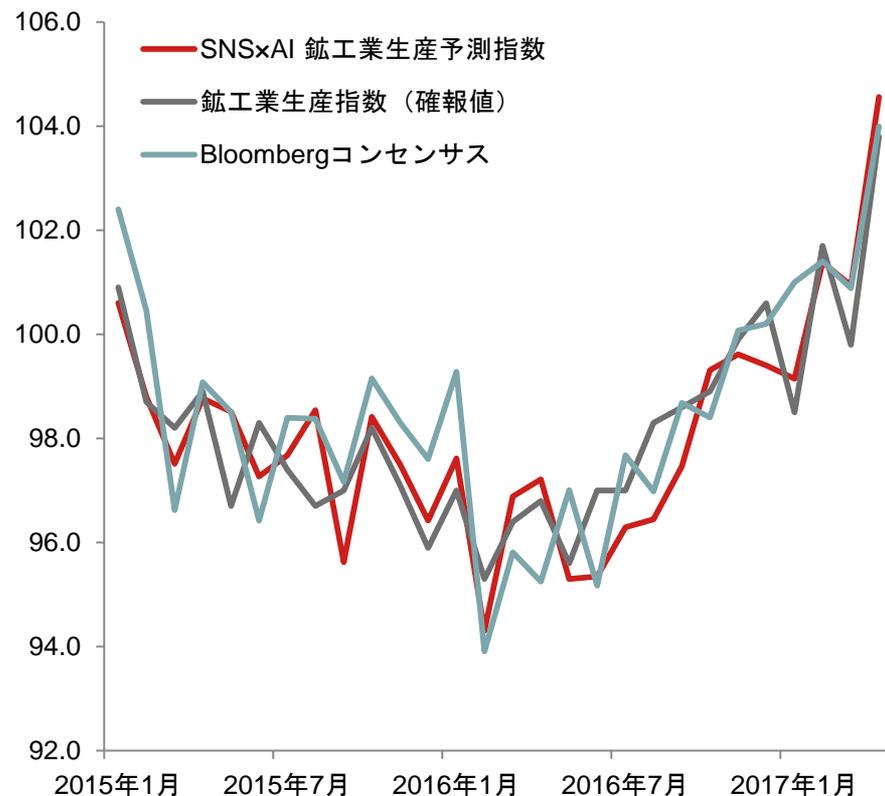


*) RMSE: 平均二乗誤差。計測期間の誤差を二乗して平均したものの平方根。

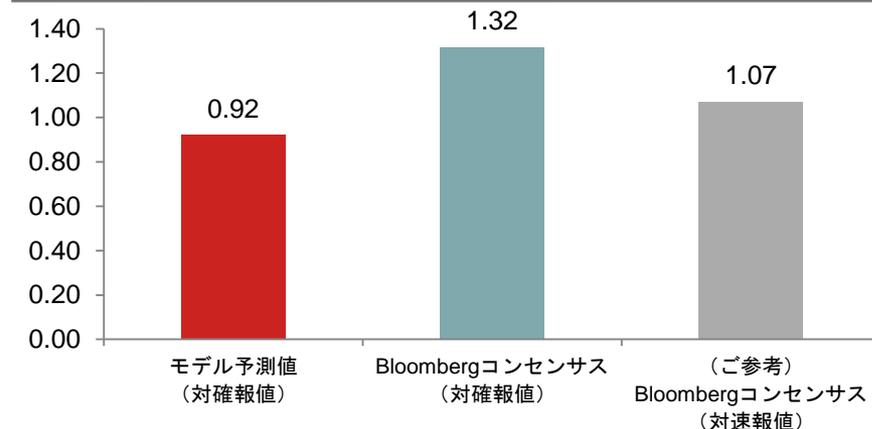
人間の専門家の予測精度との比較

確報値の予測精度を比較すると、SNSによる予測はBloombergコンセンサス予想と比較しても高い精度を達成しています。(Bloombergコンセンサス予想は確報値より速報値に対して高い精度を持っています)

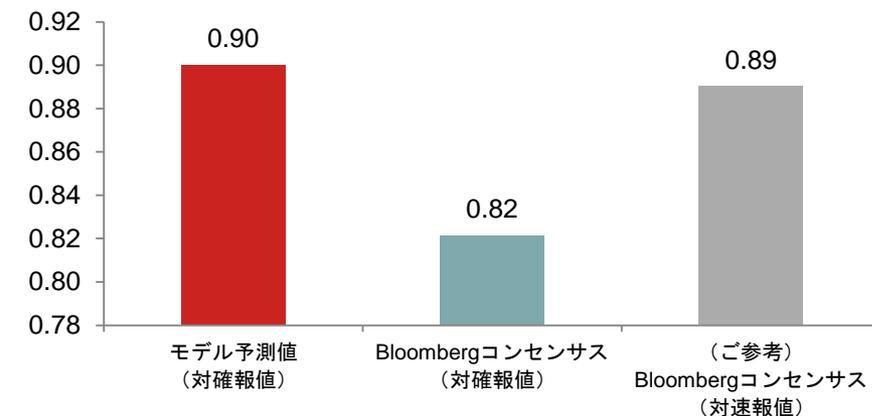
IIP実績値とモデルによる予測 / コンセンサス予想の時系列



RMSE (平均二乗誤差)



相関係数



出所) Bloomberg, NRI-IDSデータ等より野村證券作成

注) Bloombergコンセンサスの指数化では、前月確報値(年間補正前)に対して予測変化率の中央値を乗じて計算した。RMSE、相関係数の計算の「対確報値」では確報値として年間補正後の鉱工業資産指数を用いた。分析期間は2015年1月から2017年4月まで。