

کنون‌بینی رشد اقتصادی فصلی ایران با استفاده از تحلیل احساس متون خبری

* مرتضی بیرانوند

** سید محمد جواد رزمی

*** سید سعید ملک الساداتی

DOI: <https://doi.org/10.22096/esp.2025.2044951.1795>

[تاریخ دریافت: ۱۴۰۴/۰۸/۱۴ - تاریخ پذیرش: ۱۴۰۴/۱۱/۰۸]

چکیده

این پژوهش با استفاده از تحلیل احساس اخبار اقتصادی، رویکرد جدیدی را برای کنون‌بینی رشد اقتصادی سه‌ماهه ایران ارائه می‌دهد. در این راستا از اخبار روزانه اقتصادی به زبان فارسی و از ابتدای سال ۱۳۸۴ تا انتهای فصل سوم سال ۱۴۰۲ استفاده شده است. در ادامه، جهت‌گیری حسی (منفی، خنثی، مثبت) متن هر خبر و نمره حسی آن با استفاده از نسخه فارسی واژه‌نامه سنتی استرنگت مشخص و با تجمیع فصلی نمرات حسی اخبار، شاخص حسی ایجاد شده است. نتایج نشان می‌دهد که این شاخص از نظر آماری با رشد اقتصادی سه‌ماهه رابطه معنی‌داری دارد و به‌مثابه منبع اطلاعات مفید درباره روندهای کوتاه‌مدت اقتصادی در ایران عمل می‌کند؛ همچنین استفاده از این شاخص در مقایسه با مدل سری زمانی تک متغیره، بهبود معیارهای ارزیابی مدل را به همراه دارد و به درک شرایط اقتصادی در زمان واقعی و پیش از انتشار آمار رسمی کمک می‌کند.

واژگان کلیدی: اخبار اقتصادی؛ کنون‌بینی؛ رشد اقتصادی؛ تحلیل احساس.
طبقه‌بندی موضوعی: C55، C53، E17۰O47.

* دانشجوی دکتری اقتصاد، دانشکده علوم اداری و اقتصاد، دانشگاه فردوسی، مشهد، ایران. (نویسنده مسئول)

Email: beiranvand.mo@gmail.com

** استاد، دانشکده علوم اداری و اقتصاد، دانشگاه فردوسی، مشهد، ایران.

*** استادیار، دانشکده علوم اداری و اقتصاد، دانشگاه فردوسی، مشهد، ایران.

Email: mjrazmi@um.ac.ir

Email: msadati@um.ac.ir



۱. مقدمه

برای طراحی سیاست‌های اقتصادی مناسب، سیاست‌گذاران باید ارزیابی و درک دقیقی از وضعیت فعلی اقتصاد داشته باشند. سایر عوامل اقتصادی نیز نیاز به چنین درکی دارند تا بتوانند تصمیم‌های صحیح اقتصادی بگیرند. آمار رسمی که در بازه‌های زمانی مختلف به صورت عمومی منتشر می‌شوند از مهم‌ترین منابع اطلاعاتی برای ارزیابی و درک وضعیت اقتصاد محسوب می‌شوند. انتشار این آمار به دلیل مسائل مختلفی همچون دشواری‌های آمارگیری و پیچیدگی‌های محاسباتی با تأخیر صورت می‌گیرد. برای مثال در ایران، مطابق با تقویم انتشارات و تجدیدنظرهای مرکز آمار ایران، برآورد اولیه فصلی از تولید ناخالص داخلی و رشد اقتصادی به مثابه اصلی‌ترین شاخص‌های اقتصاد کلان با ۶۰ تا ۹۰ روز تأخیر منتشر می‌شود. این تأخیر باعث می‌شود که از وضعیت فعلی اقتصاد و حتی گذشته اخیر (Recent Past)، اطلاعات ناقصی در دسترس باشد. در ادبیات اقتصادی رفع مسئله عدم دسترسی به اطلاعات به هنگام از وضعیت متغیرهای اقتصادی، ذیل مفهومی تحت عنوان کنون‌بینی (Nowcasting) تعریف می‌شود. در کنون‌بینی رشد اقتصادی، هدف استفاده از هرگونه اطلاعات بالقوه مفید برای برآورد این متغیر اقتصادی قبل از انتشار آمار رسمی است. این اطلاعات طیف متنوعی از داده‌های مرسوم و کم‌بسامد مانند آمار بیکاری و تجارت گرفته تا داده‌های جدید و پر بسامد مانند تراکنش‌های پرداخت الکترونیکی یا داده‌های موتورهای جست‌وجو را دربرمی‌گیرد. در ایران، استفاده از این اطلاعات با محدودیت‌های ویژه‌ای مواجه است. از یک سو، سایر داده‌های رسمی نیز با تأخیر منتشر می‌شوند و از سوی دیگر، کلان‌داده‌هایی همچون آمار تراکنش‌های پرداخت الکترونیکی نیز به ندرت در دسترس عموم قرار دارند. به دلیل این محدودیت‌ها این پژوهش استفاده از منابع داده متون خبری را برای کنون‌بینی رشد اقتصادی پیشنهاد می‌کند. متون خبری به صورت به هنگام و برخط منتشر می‌شوند، وقفه انتشار ندارند و دسترسی به آنها محدودیت کمتری دارد. علاوه بر این، متون خبری نه تنها منبع کلیدی اطلاعات هستند، بلکه ابزاری برای ارزیابی احساسات و دیدگاه‌های روزنامه‌نگاران، تحلیلگران و کارشناسان درباره شرایط اقتصادی نیز می‌باشند. متون خبری، همانند اکثر ارتباطات انسانی، اغلب شامل محتوای روایی و احساسات هستند و به همین دلیل، ساخت شاخص‌های حسی مبتنی بر آنها می‌تواند ابزاری مستقیم برای ارزیابی انواع شوک‌های اقتصادی یا نماینده‌ای از مواردی مانند اطمینان یا انتظارات باشد. در مطالعات اقتصادی، داده‌های حسی معمولاً به منزله یکی از متغیرهای توضیحی استاندارد به

مدل‌ها اضافه می‌شوند و سپس عملکرد مدل‌های پیش‌بینی با این ادغام مورد بررسی قرار می‌گیرد.^۱ اگرچه استفاده از داده‌های متنی به زبان‌های مختلف در مطالعات اقتصادی رایج بوده است، این پژوهش از اولین مطالعاتی است که داده‌های متنی خبری فارسی را به کار می‌گیرد. نوآوری دیگر این پژوهش ساخت شاخص حسی برای اقتصاد ایران است که علاوه بر دارا بودن توانایی اندازه‌گیری احساسات اقتصادی، پیش‌بینی‌های اقتصادی را به‌ویژه در افق‌های زمانی کوتاه‌مدت بهبود می‌بخشد. برخلاف شاخص‌های مبتنی بر پرسشنامه، این شاخص قابلیت ردیابی مستمر تغییرات در احساسات و انتظارات عوامل اقتصادی را با کمترین وقفه زمانی دارد و به‌روزرسانی آن آسان است. در ادامه پژوهش، ابتدا پیشینه نظری با تأکید بر مفاهیم، تاریخچه و رویکردهای کنون‌بینی و تحلیل احساس در ادبیات اقتصادی بررسی خواهد شد. در گام بعد، پژوهش‌هایی مرور می‌شوند که از داده‌های متنی خبری و روش تحلیل احساس برای پیش‌بینی متغیرهای اقتصاد کلان استفاده کرده‌اند. در قسمت روش پژوهش، ضمن معرفی مدل استفاده‌شده در پژوهش و معیارهای ارزیابی عملکرد این مدل، داده‌های استفاده‌شده در این پژوهش معرفی می‌شوند. در انتها، بعد از تخمین مدل و با توجه به نتایج، پیشنهادهایی ارائه خواهد شد.

۲. پیشینه تحقیق

۲-۱. پیشینه نظری

۲-۱-۱. کنون‌بینی

بانبورا و همکاران،^۲ کنون‌بینی را به معنای پیش‌بینی حال، گذشته بسیار نزدیک و آینده بسیار نزدیک تعریف کرده‌اند. این اصطلاح از ترکیب دو واژه اکنون (Now) و پیش‌بینی (Forecasting) ساخته شده است. اصطلاح کنون‌بینی نخستین بار در حوزه هواشناسی و در پاسخ به نیاز روزافزون بخش صنعت و خدمات به پیش‌بینی‌های به‌هنگام از شرایط جوی با استفاده از آخرین اطلاعات در دسترس به کار رفته است. این اصطلاح بعد از مطالعه جیانونی

1. Andres Algaba et al., "Econometrics meets sentiment: An overview of methodology and applications," *Journal of Economic Surveys* 34, no. 3 (2020): 517-518.

2. Marta Banbura, Domenico Giannone, and Lucrezia Reichlin, "Nowcasting," in *The Oxford handbook of economic forecasting*, eds. Clements, Michael P. and David F. Hendry (OUP USA, 2011), 193.

و همکاران^۳ با عنوان «کنون‌بینی تولید ناخالص داخلی و تورم: محتوای اطلاعاتی بلادرنگ داده‌های اقتصاد کلان» در اقتصاد رایج شد و بعد از آن انبوهی از مقالات منتشر شدند که هدفشان ارائه برآوردهای به‌هنگام از شاخص‌های اقتصاد کلان به‌ویژه تولید ناخالص داخلی بوده است.

دلیل اقبال اقتصاددانان به این موضوع، تأخیر قابل توجه در انتشار اطلاعات شاخص‌های اقتصادی است که منجر به عدم شناسایی تحولات اقتصادی در زمان واقعی می‌شود. برای مثال در کشورهای G7، اولین برآورد رسمی تولید ناخالص داخلی فصلی با تأخیری در حدود یک تا دو ماه بعد از سه‌ماهه مرجع منتشر می‌شود.^۴ اهمیت روزافزون در دسترس بودن برآوردهای به‌هنگام از شرایط اقتصادی به‌ویژه پس از بحران مالی سال ۲۰۰۸، پژوهشگران و متخصصان اقتصادی را به توسعه مدل‌های اقتصادسنجی سوق داده است که به‌طور خاص هدفشان پیش‌بینی متغیرهای اقتصاد کلان در دوره فعلی یا حداکثر دوره آینده است.^۵ این مدل‌ها اغلب بر مبنای شاخص‌های با تأثیر بالا و وقفه انتشار اندک توسعه یافته‌اند. این شاخص‌ها به دو دسته شاخص‌های سخت و نرم تقسیم می‌شوند. شاخص‌های سخت مانند میزان تولیدات صنعتی، شاخص‌های خرده‌فروشی و نرخ بیکاری که به‌طور مستقیم اطلاعات کمی از سطح فعالیت اقتصادی را نشان می‌دهند. از طرف دیگر، شاخص‌های نرم مانند شاخص مدیران خرید بر اساس نظرسنجی‌ها و پرسشنامه‌ها، احساسات و انتظارات عوامل اقتصادی را بازتاب می‌دهند.^۶ اخیراً استفاده از مجموعه داده‌های بزرگ که از منابع گوناگون اطلاعاتی ناشی می‌شوند در زمینه اقتصاد به‌طور گسترده‌ای فراگیر شده است. این کلان‌داده‌ها امکان دسترسی به اطلاعات ارزشمندی را قبل از انتشار آمار رسمی فراهم می‌کنند و تحلیل‌هایی نوآورانه و دقیق‌تر در زمینه‌های مختلف اقتصادی را ممکن می‌سازند.^۷ این مجموعه داده‌های بزرگ، دامنه گسترده‌ای از کلان‌داده‌ها همچون مصرف برق و انرژی،

3. Domenico Giannone, Lucrezia Reichlin, and David H. Small, "Nowcasting GDP and inflation: the real-time informational content of macroeconomic data releases," *ECB working paper*, no. 633 (2006): 1-52.

4. Danilo Cascaldi-Garcia, Matteo Luciani, and Michele Modugno, "Lessons from Nowcasting GDP across the World," *International Finance Discussion Paper* 1385 (2023): 5-6.

5. Boriss Siliverstovs, "New York FED staff nowcasts and reality: What can we learn about the future, the present, and the past?," *Econometrics* 9, no. 1 (2021): 1.

6. Mary Keeney, Bernard Kennedy, and Joelle Liebermann, "The value of hard and soft data for short-term forecasting of GDP," *Economic Letters*, no. 11 (2012): 1.

7. Ziqi Zang, "Three Essays on Big Data in International Finance" (PhD diss., UCLA, 2019), 1.

داده‌های ترافیکی و حمل‌ونقل، داده‌های موتورهای جست‌وجو، داده‌های پرداخت الکترونیکی، تصاویر ماهواره‌ای و اسناد متنی را دربرمی‌گیرد.^۸ داده‌های متنی در جایگاه یکی از مهم‌ترین انواع کلان‌داده‌ها، به دلیل ماهیت بدون‌ساختار خود با داده‌های سنتی متفاوت هستند. آنها منبعی ارزشمند از اطلاعات کیفی در زمینه اقتصاد هستند که اغلب از طریق قوانین پیچیده زبانی رمزگذاری شده‌اند.^۹

داده‌های متنی از منابع گوناگونی مانند اسناد دیجیتال، خبرگزاری‌ها، رسانه‌های اجتماعی، نظرات آنلاین، وبلاگ‌ها و منابع دیگر در دسترس قرار دارند. شناسایی و استخراج الگوهای قابل فهم از این داده‌ها یکی از اهداف اصلی در حوزه متن‌کاوی است. این حوزه از زمینه داده‌کاوی الهام گرفته است.^{۱۰} در مطالعات اقتصادی، اغلب با استفاده از تحلیل احساس متون به مثابه یکی از کاربردهای مهم متن‌کاوی، احساس و جهت‌گیری متن کشف و شاخص‌های حسی بر مبنای آن ساخته می‌شود، سپس این شاخص‌ها به صورت مستقل یا در ترکیب با سایر اطلاعات در مدل‌های پیش‌بینی اقتصادی به کار گرفته می‌شوند.

۲-۱-۲. تحلیل احساس

تحلیل احساس، که به آن عقیده‌کاوی (Opinion Mining) نیز می‌گویند، حوزه مطالعاتی است که نظرات، احساسات، ارزیابی‌ها، نگرش‌ها و عواطف افراد را نسبت به موجودیت‌هایی از قبیل محصولات، خدمات، سازمان‌ها، افراد، مسائل، رویدادها، موضوعات و ویژگی‌های هر یک از آنها تجزیه و تحلیل می‌کند.^{۱۱} تحلیل احساس به منزله شاخه‌ای از یادگیری ماشین، داده‌کاوی، پردازش زبان طبیعی (Natural Language Processing) و زبان‌شناسی رایانشی (Computational Linguistics) در نظر گرفته می‌شود که عناصری از جامعه‌شناسی و

8. Luca Barbaglia et al., "Testing big data in a big crisis: Nowcasting under COVID-19," *International Journal of Forecasting* 39, no. 4 (2023): 1548-1563; Lucas Manchado Marcos, "Nowcasting with Alternative Data" (Bachelor's thesis, Polytechnic University of Catalonia, 2023), 8.

9. Valentina Aprigliano et al., "The power of text-based indicators in forecasting Italian economic activity," *International Journal of Forecasting* 39, no. 2 (2023): 791.

10. Andreas Hotho, Andreas Nürnberger, and Gerhard Paaß, "A brief survey of text mining," *Journal for Language Technology and Computational Linguistics* 20, no. 1 (2005): 19-62.

11. Bing Liu, *Sentiment analysis and opinion mining* ([n.p]: Springer Nature, 2022), 7.

روانشناسی را نیز در خود جای داده است.^{۱۲}

برخلاف زبان‌شناسی و پردازش زبان طبیعی که پیشینه پژوهشی طولانی دارند، اصطلاح تحلیل احساس احتمالاً اولین بار در ابتدای قرن بیست و یکم در پژوهش‌ها پدیدار شده است، ولی امروزه تحلیل احساس تحت سه عامل اصلی به یک زمینه پژوهشی بسیار فعال تبدیل شده است. اولین عامل در گسترش آن، کاربرد فراوان در حوزه‌های مختلف است. برای مثال، کاربردهای تجاری تحلیل احساس در بخش صنعت، محرکی قوی برای تحقیقات بیشتر در این زمینه است. دوم، گسترش دامنه استفاده از آن باعث بروز مسائل فنی جدیدی شده است که حل آن نیازمند پژوهش‌های جدید و نوآورانه هستند. سومین و احتمالاً مهم‌ترین عامل این است که برای نخستین بار در تاریخ بشر، مقدار عظیمی از داده‌ها در اینترنت و شبکه‌های اجتماعی تولید و منتشر می‌شوند که شامل نظرات و احساسات افراد هستند. بدون این مجموعه داده‌های بزرگ، انجام دادن پژوهش‌های زیادی امکان‌پذیر نبود. بنابراین، نباید عجیب باشد که آغاز و رشد سریع تحلیل احساس همزمان با شکوفایی رسانه‌های اجتماعی رخ داده است.^{۱۳} اگرچه تحلیل احساس برای کشف احساس در منابع اطلاعاتی مانند صوت، عکس و ویدیو نیز استفاده می‌شود، اما اغلب پژوهش‌ها متمرکز بر کشف احساسات در متون طبیعی بوده‌اند.^{۱۴}

۱-۲-۱-۲. وظایف تحلیل احساس

می‌توان وظایف اصلی و مسائل اساسی را که تحلیل احساس با آنها روبه‌رو است در سه دسته تشخیص ذهنیت (Subjectivity Detection)، تشخیص قطبیت (Polarity Detection) و وظایف تکمیلی تقسیم کرد.^{۱۵}

۱. تشخیص ذهنیت: هدف تشخیص ذهنیت، طبقه‌بندی متن به صورت خودکار به

12. Lin Yue et al., "A survey of sentiment analysis in social media," *Knowledge and Information Systems* 60 (2019): 618.

13. Liu. *Sentiment analysis*, 7-8.

۱۴. عبدالصمد کرامت‌فر، «مدل‌سازی چند جریانی زمینه نظرات برای تحلیل احساس» (رساله دکترا، دانشگاه قم، ۱۴۰۰)، ۳.

15. Akshi Kumar and Mary Sebastian Teeja, "Sentiment analysis: A perspective on its past, present and future," *International Journal of Intelligent Systems and Applications* 4, no. 10 (2012): 4;

زینب رجبی، محمدرضا ولوی، و مریم حورعلی، «مروری بر روش‌های تحلیل احساس در متون فارسی»، فصل‌نامه پردازش علائم و داده‌ها ۱۹، شماره ۲ (تابستان ۱۴۰۱): ۱۱۲.

جملات ذهنی یا حاوی نظر در مقابل جملات عینی یا خنثی است. از این وظیفه، تحلیل احساس برای شناسایی جملات ذهنی یعنی جملاتی که بیانگر نظر، دیدگاه یا احساس نویسنده است و حذف جملات عینی که تنها واقعیتی را بیان می‌کند و نظر و حسی در آنها نیست، استفاده می‌شود.

۲. تشخیص قطبیت: تشخیص قطبیت به‌مثابه اصلی‌ترین و معروف‌ترین وظیفه در تحلیل احساس شناخته می‌شود به‌طوری که در بسیاری از پژوهش‌ها این دو عبارت به‌جای یکدیگر استفاده می‌شوند. این وظیفه می‌تواند یک طبقه‌بندی باینری (مثبت یا منفی) یا طبقه‌بندی چندگانه (بسیار منفی، منفی، خنثی، مثبت یا بسیار مثبت) ارائه دهد.

۳. وظایف تکمیلی: ازجمله این وظایف می‌توان به ایجاد منابع واژگان حسی (Sentiment Lexicon)، استخراج جنبه‌ها یا ویژگی‌های یک محصول یا خدمت از متن، تشخیص نظرات جعلی و تشخیص کنایه را نام برد.

۲-۲-۱-۲. سطوح تحلیل احساس

تحلیل احساس متون طبیعی را می‌توان در سه سطح شامل جمله، سند و جنبه انجام داد.

۱. تحلیل احساس در سطح جمله: هدف از طبقه‌بندی احساسات در سطح جمله، ارزیابی اسناد بر اساس نظر موجود در هر جمله و طبقه‌بندی جملات بر حسب دارا بودن احساسات مثبت، منفی یا خنثی است.^{۱۶}

۲. تحلیل احساس در سطح سند: در این سطح، هدف طبقه‌بندی قطبیت کلی سند است به این معنا که آیا یک سند در کل، یک احساس یا نظر منفی یا مثبت را بیان می‌کند. هر سند براساس احساس کلی نویسنده در مورد یک موجودیت واحد (برای مثال، محصولی مشخص) طبقه‌بندی می‌شود. طبقه‌بندی بهترین عملکرد را در سطح سند زمانی دارد که سند توسط شخصی نوشته شده باشد و برای اسنادی که چند موجودیت را ارزیابی یا مقایسه می‌کنند، مناسب نیست.^{۱۷}

۳. تحلیل احساس در سطح جنبه: اگرچه تحلیل احساس متن در سطح سند و جمله در

16. Liu. *Sentiment analysis*, 49.

17. Marouane Birjali, Mohammed Kasri, and Abderrahim Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Systems* 226, no. 107134 (2021): 3.

بسیاری از موارد مفید است، اما تمام جزئیات لازم را ارائه نمی‌دهد. برای مثال، اینکه یک سند احساس مثبتی درباره یک موجودیت خاص دارد، به این معنی نیست که نویسنده به همه جنبه‌های آن موجودیت احساس مثبت دارد و به‌طور مشابه، اینکه یک سند احساس منفی داشته باشد، نشان‌دهنده نظر منفی نویسنده درباره همه جنبه‌های آن موجودیت نیست. هدف تحلیل احساس در این سطح، شناسایی و استخراج جنبه‌های مختلف از متن و سپس تعیین قطبیت هر جنبه است.^{۱۸}

۳-۲-۱-۲. رویکردهای تحلیل احساس

تحلیل احساس اغلب با سه رویکرد یادگیری ماشین (Approach Machine Learning)، رویکرد مبتنی بر واژه‌نامه (Approach Lexicon-based) و رویکرد ترکیبی (Approach Hybrid) انجام می‌شود.^{۱۹}

۱. رویکرد یادگیری ماشین: در رویکرد یادگیری ماشین برای طبقه‌بندی قطبیت احساسات، از مجموعه داده‌های آموزش و آزمایش استفاده می‌شود. این رویکرد به یکی از سه شیوه یادگیری با نظارت (Supervised Learning)، یادگیری بدون نظارت (Unsupervised Learning) و یادگیری نیمه نظارتی (Semi-supervised Learning) انجام می‌شود.^{۲۰}

۲. رویکرد مبتنی بر واژه‌نامه: رویکرد مبتنی بر واژه‌نامه یا واژه‌نامه‌محور یکی از مهم‌ترین رویکردهای تحلیل احساس است. در این روش، سیستم‌های توسعه‌یافته با استفاده از واژه‌نامه‌های از پیش ساخته که متشکل از واژگان و عبارات حاوی احساس و نمره عددی احساس آنها است، یک نمره مجموع برای عبارت، جمله یا سند موردنظر را محاسبه می‌کنند و همین عدد را به عنوان احساس آن در نظر می‌گیرند.^{۲۱}

۳. رویکرد ترکیبی: این رویکرد، توان عملیاتی تحلیل واژگانی را با انعطاف‌پذیری رویکردهای یادگیری ماشین ترکیب می‌کند، به‌طوری که می‌تواند از مزایای هر دو رویکرد بهره‌مند شود و در عین حال از محدودیت‌های آنها جلوگیری کند. ویژگی اصلی این رویکرد

18. Liu. *Sentiment analysis*, 11.

19. Birjali, Kasri, and Beni-Hssane, "A comprehensive survey," 8-10.

20. Zohreh Madhoushi, Abdul Razak Hamdan, and Suhaila Zainudin, "Sentiment analysis techniques in recent works," *science and information conference (SAI)* (2015): 288-291.

۲۱. کرامت‌فر، «مدل‌سازی چند جریانی»، ۴.

به ارث بردن دقت بالا از یادگیری ماشین و ثبات از رویکرد بر مبتنی بر واژه‌نامه است.^{۲۲}

۲-۲. پیشنهاد تجربی

همگام با گسترش انواع داده‌های متنی و پیشرفت روش‌های ذخیره‌سازی و تجزیه و تحلیل این داده‌ها، مطالعات اقتصادی زیادی بر استفاده از این نوع کلان‌داده متمرکز شده‌اند. گارز^{۲۳} با مرور مطالعات اقتصادی که از داده‌های متنی استفاده کرده‌اند؛ برخی از این مطالعات را برحسب نوع رسانه، هدف مطالعه و روش تحلیل دسته‌بندی کرده است. در این بخش، برخی پژوهش‌هایی که از داده‌های متنی خبری و به‌کارگیری روش‌های مختلف متن‌کاوی از جمله تحلیل احساس برای پیش‌بینی متغیرهای اقتصاد کلان استفاده کرده‌اند، اشاره می‌شود. در مطالعه سئو و همکاران^{۲۴} از یک مدل انتقالی یادگیری عمیق برای ساختن یک شاخص احساس مبتنی بر اخبار اقتصادی کره جنوبی استفاده شده است. این شاخص، که از اطلاعات روزانه اخبار بهره می‌برد، قادر است تغییرات در احساسات اقتصادی را قبل از انتشار شاخص‌های رسمی ماهیانه‌ای مانند شاخص احساس مصرف‌کننده شناسایی کند و به‌مثابه یک ابزار پیشرو در تشخیص نقاط عطف و علل نوسانات اقتصادی عمل کند.

در پژوهشی دیگر ابریکلیانو و همکاران^{۲۵} نشان می‌دهند که با استفاده از مجموعه‌ای از شاخص‌های حسی که از مقالات خبری استخراج می‌شوند، می‌توان پیش‌بینی‌های دقیق‌تری درباره شاخص‌های اقتصاد کلان ایتالیا از جمله رشد تولید ناخالص داخلی و تولید صنعتی داشت. این شاخص‌های حسی، اطلاعات به‌هنگامی در مورد وضعیت اقتصاد ارائه می‌دهند و قادرند نقاط عطف و روندهای اقتصادی را سریع‌تر از شاخص‌های مرسوم شناسایی کنند و به همین دلیل در پیش‌بینی‌های کوتاه‌مدت به‌ویژه در دوره‌های افزایش عدم اطمینان، کارآمدتر هستند؛ همچنین نتایج نشان داد که ترکیب این شاخص‌های حسی با منابع داده دیگر، دقت مدل‌های پیش‌بینی را بهبود می‌بخشد. شیپرو و همکاران^{۲۶} با استفاده از رویکرد مبتنی بر

22. Birjali, Kasri, and Beni-Hssane, "A comprehensive survey," 16-17.

23. Marcel Garz, "Macroeconomic news: A literature survey and methodological guidelines," SSRN 4643970 (2023):1-30.

24. Beomseok Seo, Younghwan Lee, and Hyungbae Cho, "Measuring News Sentiment of Korea Using Transformer," *Korean Economic Review* 40, no. 1 (2024): 149-176.

25. Valentina Aprigliano et al., "The power of text-based," 791-808.

26. Adam Hale Shapiro, Moritz Sudhof, and Daniel J. Wilson, "Measuring news sentiment," *Journal of econometrics* 228, no. 2 (2022): 221-243.

واژه‌نامه، سری‌های زمانی حسی را با تحلیل احساس اخبار اقتصادی در دوره زمانی ۱۹۸۰-۲۰۱۵ ایجاد کردند و با استفاده از روش‌های یادگیری ماشین نشان دادند که این سری‌های زمانی می‌توانند برای پیش‌بینی به‌هنگام متغیرهای اقتصادی در آمریکا به کار گرفته شوند. سکی و همکاران^{۲۷} یک شاخص حسی بر اساس مقالات خبری برای ژاپن ساختند که همبستگی قوی با شاخص‌های رسمی مبتنی بر نظرسنجی دارند و در عین حال ینش بیشتری در مورد روند اقتصاد کلان و احساسات فعالان اقتصادی منعکس می‌کند. مطالعه تیلی و همکاران^{۲۸} نشان داد که با استفاده از شبکه‌های عصبی و جداسازی سیگنال‌های مهم از حجم زیادی از داده‌های اخبار بین‌المللی، می‌توان از داده‌های حسی استخراج‌شده از این اخبار بهره‌برداری کرد. به‌کارگیری داده‌های حسی مستخرج از این اخبار می‌تواند دقت مدل‌های پیش‌بینی اقتصاد کلان را به‌طور قابل‌توجهی در مقایسه با مدل‌های خودرگرسیون و رویکردهای مرسوم که صرفاً بر داده‌های اقتصادی یا معیارهای احساسی ساده متکی هستند، افزایش دهد. رمبوکاسینگ و کایتکوفسکی^{۲۹} با استفاده از رویکرد مبتنی بر واژه‌نامه و الگوریتم‌های یادگیری ماشینی، رابطه بین احساسات مستخرج از مقالات خبری روزنامه‌های بزرگ بریتانیا و شاخص‌های اقتصادی مختلف مانند تورم، رشد تولید ناخالص داخلی و نرخ بیکاری را بررسی کردند. نتایج نشان داد که روایت‌های گزارش‌شده در رسانه‌ها می‌توانند اطلاعاتی پیشرو برای متغیرهای اقتصاد کلان فراهم کنند و استفاده از این اطلاعات در مقایسه با مدل‌های استاندارد، بهبود قابل‌توجهی در عملکرد مدل‌های پیش‌بینی بیکاری و رشد تولید ناخالص داخلی داشته است. با این حال، قدرت پیش‌بینی آنها برای شاخص‌های اقتصادی مختلف، متفاوت است. ثورسروود^{۳۰} توانایی استفاده از داده‌های متنی اخبار را به‌مثابه مکمل شاخص‌های اقتصادی استاندارد برای نظارت و درک پویایی چرخه تجاری اثبات کرد. این شاخص مبتنی بر اخبار که با استفاده از روش‌های پردازش زبان طبیعی از مجموعه بزرگی از مقالات خبری نروژ ایجاد شده است در تواتر بالاتر (روزانه) و بدون تأخیر در دسترس است

27. Kazuhiro Seki, Yusuke Ikuta, and Yoichi Matsubayashi, "News-based business sentiment and its properties as an economic index," *Information Processing & Management* 59, no. 2, 102795 (2022): 1-16.

28. Sonja Tilly, Markus Ebner, and Giacomo Livan, "Macroeconomic forecasting through news, emotions and narrative," *Expert Systems with Applications* 175, no. 114760 (2021): 1-12.

29. Dooruj Rambaccussing and Andrzej Kwiatkowski, "Forecasting with news sentiment: Evidence with UK newspapers," *International Journal of Forecasting* 36, no. 4 (2020): 1501-1516.

30. Leif Anders Thorsrud, "Words are the new numbers: A newsy coincident index of the business cycle," *Journal of Business & Economic Statistics* 38, no. 2 (2020): 393-409.

و می‌تواند روندهای اقتصادی و نقاط عطف را سریع‌تر به تصویر بکشد. آردیا و همکاران^{۳۱} نشان دادند که افزودن داده‌های حسی مبتنی بر اخبار به مدل‌های پیش‌بینی در مقایسه با استفاده از روش‌هایی که صرفاً بر شاخص‌های اقتصادی و مالی تکیه دارند، باعث افزایش دقت پیش‌بینی نرخ رشد نه‌ماهه و سالانه تولید صنعتی آمریکا می‌شود. فیوریگل و گوردن^{۳۲} در کنار روش شمارش واژگان، رویکرد مدل‌سازی معنایی را توسعه دادند که با تحلیل احساس اخبار مالی می‌تواند شاخص‌های اقتصاد کلان را برای آلمان و اتحادیه اروپا به‌طور مؤثر پیش‌بینی کند.

۳. روش تحقیق

۳-۱. تحلیل احساس با رویکرد واژه‌نامه محور

واژه‌نامه‌های زیادی برای تعیین قطبیت کلمه‌ها وجود دارند. هر واژه‌نامه، سازوکار خاصی را برای تعیین قطبیت استفاده می‌کند و روش خاصی را برای نشان دادن قطبیت کلمه‌ها به کار می‌برد. برای مثال، کلماتی مانند «خوب»، «زیبا» و «شگفت‌انگیز» احساسات مثبتی را به انسان منتقل می‌کنند، در حالی که کلماتی همچون «بد»، «زشت» و «ترسناک» دارای قطبیت منفی هستند. این نکته مهم است که بیشتر کلمات دارای بار حسی در واقع صفت و قید هستند، اما برخی نام‌ها مانند «زباله» و افعالی مثل «نفرت داشتن» و «عشق ورزیدن» نیز می‌توانند دارای بار حسی باشند. این کلمات ممکن است دارای وزن یا بدون وزن باشند. وزن در اینجا به معنای یک عدد یا احتمال است که برای هر کلمه در نظر گرفته می‌شود تا سطح مثبت یا منفی بودن آن کلمه را نشان دهد. این رویکرد اغلب برای محاسبه جهت‌گیری اسناد با توجه به جهت‌گیری معنایی کلمات و عبارات درون اسناد استفاده می‌شود.^{۳۳} سنتی‌استرنگت (Strength Senti) از جمله این واژه‌نامه حسی است که توسط تلوال و

31. David Ardia, Keven Bluteau, and Kris Boudt, "Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values," *International Journal of Forecasting* 35, no. 4 (2019): 1370-1386.

32. Stefan Feuerriegel and Julius Gordon, "News-based forecasts of macroeconomic indicators: A semantic path model for interpretable predictions," *European Journal of Operational Research* 272, no. 1 (2019): 162-175.

33. Fatemeh Hemmatian and Mohammad Karim Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis," *Artificial intelligence review* 52, no. 3 (2019): 1526-1531.

همکاران^{۳۴} برای زبان انگلیسی و در ادامه برای سایر زبان‌ها از جمله زبان فارسی، توسعه پیدا کرده است و برای تحلیل احساس و محاسبه نمره حسی متون از لیست‌های زیر استفاده می‌کند:

۱. لیست کلمات حسی: مجموعه‌ای از اصطلاحات مثبت و منفی است که هرکدام دارای مقداری از ۵- تا ۵ هستند.

۲. لیست کلمات پرسشی: حاوی چند کلمه پرسشی پرتکرار است.

۳. لیست کلمات نفی: حاوی چند کلمه نفی پرتکرار برای معکوس کردن کلمات حسی بعد از خود است.

۴. لیست نمادهای حسی: مکمل لیست اول و ترکیب علامت‌هایی است که برای بیان احساسات استفاده شد.

برخی کلمات نمونه از این فهرست‌ها در جدول (۱) نشان داده شده است.

جدول (۱): نمونه‌ای از لیست واژه‌های سنتی استرنگت

نوع لیست	نمونه کلمات	نمره حسی
لیست کلمات حسی	ارزنده	+۲
	عالی	+۳
	تحمیل	-۲
	تنش	-۳
لیست کلمات پرسشی	چرا	-
	چگونه	-
لیست کلمات نفی	هرگز	-
	هیچ	-
لیست نمادهای حسی	:D	+۱
	:('	-۱

34. Mike Thelwall et al., "Sentiment strength detection in short informal text," *Journal of the American society for information science and technology* 61, no. 12 (2010): 2544-2558.

کنون پینی رشد اقتصادی فصلی ایران با استفاده از ... / پیرانوند، رزمی و ... ۹۵

سنتی استرنگت، احساس کلی یک متن طولانی را با احساسات مثبت و منفی کلی هر جمله محاسبه می‌کند و مانند سایر الگوریتم‌هایی که تحلیل احساس بر مبنای واژه‌نامه را به کار می‌برند بر این منطق استوار است که نویسنده‌ای که احساس منفی نسبت به یک موضوع دارد، برآیند کلماتی که استفاده می‌کند منفی‌تر است و بالعکس. در مطالعات اقتصادی، این واژه‌نامه برای کشف احساس متون در زمینه‌های مختلفی به کار رفته است.^{۳۵} استفاده از نسخه فارسی این واژه‌نامه در مطالعات اقتصادی به زبان فارسی نیز سابقه دارد.^{۳۶}

۲-۳. مدل اقتصادسنجی و معیارهای ارزیابی

مدل مورد استفاده در این پژوهش مطابق رابطه (۱) قابل نمایش است که در آن Y_t رشد اقتصادی و NSI_t شاخص حساسی اخبار (News Sentiment Index) است.

$$1) Y_t = Y_{t-1} + NSI_t + u_t$$

برای اندازه‌گیری دقت پیش‌بینی مدل از معیارهای ارزیابی میانگین قدر مطلق خطا (Mean Absolute Error) (MAE)، جذر میانگین مربعات خطا (Root Mean Square Error) (RMSE) و ضریب نابرابری تایل (Theil Inequality Coefficient) (TIC) استفاده خواهد شد. هر سه معیار، مطابق با روابط (۲-۴)، مقدار خطا را از مقایسه مقادیر واقعی (Y_t) و پیش‌بینی (Y_t^f) محاسبه می‌کنند.^{۳۷}

35. Joanna Strycharz, Nadine Strauss, and Damian Trilling, "The role of media coverage in explaining stock market fluctuations: Insights for strategic financial communication," *International Journal of Strategic Communication* 12, no. 1 (2018): 67-85; Tiran Rothman, and Chen Yakar, "Empirical Analysis Towards the Effect of Social Media on Cryptocurrency Price and Volume," *European Scientific Journal, ESJ* 15 (2019): 31-52; Anna Brosius, Erika J. van Elsas, and Claes H. de Vreese, "Bad news, declining trust? Effects of exposure to economic news on trust in the European Union," *International Journal of Public Opinion Research* 32, no. 2 (2020): 223-242.

۳۶. مرتضی آهنگری آهنگرکلایی، علی سبطی، و مهدی یعقوبی، «ساخت واژگان به صورت خودکار برای تحلیل نظرات در حوزه بورس»، فصل‌نامه پردازش علائم و داده‌ها، ۲۰، شماره ۲ (تابستان ۱۴۰۲): ۳-۲۰؛ عظیم زارعی، داود فیض، و غزاله طاهری، «ارائه چارچوب هوشمندی بازار اجتماعی مبتنی بر وب ۲/۰ با استفاده از تکنیک متن‌کاوی در وبسایت‌های رسانه‌های اجتماعی (مورد مطالعه: تحلیل رقابتی در بین برندهای سامسونگ و امرسان)»، پژوهش‌های مدیریت در ایران، ۲۴، شماره ۴ (۱۳۹۹): ۹۸-۱۲۵.

۳۷. علی سوری، اقتصادسنجی پیشرفته، چاپ ۱، جلد ۲ (تهران: انتشارات نور علم، ۱۴۰۰)، ۸۷۰-۸۷۲.

$$2) MAE = \frac{\sum_{t=T+1}^{T+m} |Y_t^f - Y_t|}{m}$$

$$3) RMSE = \sqrt{\frac{\sum_{t=T+1}^{T+m} (Y_t^f - Y_t)^2}{m}}$$

$$4) TIC = \frac{\sqrt{\frac{\sum_{t=T+1}^{T+m} (Y_t^f - Y_t)^2}{m}}}{\sqrt{\frac{\sum_{t=T+1}^{T+m} (Y_t^f)^2}{m}} + \sqrt{\frac{\sum_{t=T+1}^{T+m} (Y_t)^2}{m}}}$$

m طول دوره پیش‌بینی است که از T+1 تا T+m می‌باشد.

۳-۳. مجموعه داده‌ها

در این مطالعه از داده‌های رشد اقتصادی فصلی ایران مستخرج از سایت مرکز آمار ایران به قیمت‌های ثابت سال ۱۳۹۰ و در بازه زمانی فصل اول سال ۱۳۸۴ تا فصل سوم سال ۱۴۰۲، استفاده شده است. برای ساخت شاخص حسی نیز از اخبار اقتصادی خبرگزاری فارس در همان دوره استفاده شده است. در مقدمات انجام دادن این پژوهش برای استخراج اخبار اقتصادی فارسی، خبرگزاری‌ها و وبگاه‌های خبری مختلفی بررسی و در نهایت خبرگزاری فارس برای منبع اطلاعاتی این پژوهش انتخاب شد. این انتخاب با توجه به سه معیار وجود آرشیو اخبار در دوره‌های زمانی کافی، مرجع بودن خبرگزاری و قابلیت دسترسی و استخراج اخبار به صورت برخط انجام شده است. برای برازش صحیح مدل‌های رگرسیونی باید تعداد مشاهدات به حد کافی بالا باشد به همین دلیل برخی خبرگزاری‌ها به علت عدم وجود دوره‌های زمانی کافی از آرشیو اخبار، کنار گذاشته شدند. لازم به ذکر است امکان ترکیب اخبار چند خبرگزاری برای رفع این نقص نیز بررسی شد که به علت قالب‌های متفاوت و فرمت‌های مختلف انتشار و ذخیره‌سازی اخبار هر یک از این خبرگزاری‌ها و دشواری‌های پردازشی تجمیع متون فارسی از منابع مختلف، امکان تجمیع و پردازش متون فراهم نشد. معیار دوم یعنی مرجع بودن یک خبرگزاری، تضمینی برای این نکته است که خبرگزاری برای پوشش تمامی مخاطبان و حفظ آمار بالایی بازدید از همه بخش‌های اقتصاد و از جنبه‌های مختلف خبر منتشر می‌کند و به همین دلیل احتمال مغفول ماندن یک بخش یا حوزه اقتصادی کاهش می‌یابد. از طرف دیگر، مرجع بودن یک خبرگزاری، اگرچه تضمین نمی‌کند که

سوگیری‌های احتمالی در نحوه خبررسانی کاهش یابد؛ اما احتمال انتشار اخبار جعلی و بدون منبع را کمتر می‌کند. در این پژوهش مرجع بودن خبرگزاری با توجه به معیارهای تحلیل دامنه و رتبه‌بندی سایت آکسا (Alexa) تعیین شده است. براساس این رتبه‌بندی، خبرگزاری فارس یا توجه به ترافیک ورودی سایت در سال ۱۴۰۰ رتبه ۴۲ را در میان همه سایت‌های فارسی‌زبان داشته است و در سال ۱۳۹۹ در مقایسه با سه خبرگزاری پربازدید مهر، ایسنا و ایرنا در رتبه دوم قرار گرفته است. آخرین عامل و شاید مهم‌ترین عامل در انتخاب خبرگزاری مناسب این بوده است که از لحاظ فنی امکان استخراج اخبار با روش مورد استفاده در این پژوهش وجود داشته باشد. برخی خبرگزاری‌هایی که دو معیار قبلی را داشتند با استفاده از روش‌های فنی خاص، محدودیت پاسخگویی به درخواست‌های خودکار وب کراولینگ را برای ذخیره‌سازی خودکار اخبار ایجاد کرده‌اند و خبرگزاری فارس تنها خبرگزاری بود که امکان استخراج اخبار آن با ابزار در دسترس این پژوهش فراهم شد.

۴. برآورد مدل

در ابتدا با روش خزش (Web Crawling)، ۳۰۱،۴۹۸ خبر اقتصادی در دوره زمانی مورد بررسی از پایگاه اینترنتی خبرگزاری فارس استخراج شد. در ادامه با به‌کارگیری واژه‌نامه سنتی‌استرنگت در محیط زبان برنامه‌نویسی پایتون، نمره حسی هر خبر مشخص شد. همانطور که پیش‌تر اشاره شد این ابزار از واژه‌نامه‌ای از کلمات حسی موزون استفاده می‌کند که به‌صورت دستی توسط متخصصان زبان‌شناسی انتخاب شده است. در تحلیل احساس هر خبر از همه اجزای آن خبر شامل عنوان، خلاصه و متن اصلی استفاده شده است. وقتی ابزار سنتی‌استرنگت، متنی را می‌خواند آن را به کلمات تشکیل‌دهنده آن تقسیم و علائم نگارشی را از هم جدا می‌کند. سپس، هر کلمه از نظر مطابقت با هر یک از واژه‌های حسی واژه‌نامه بررسی می‌شود و اگر مطابقت یافت شود، امتیاز حسی مرتبط حفظ می‌شود. نمره کلی برای یک جمله بالاترین نمره مثبت و منفی برای کلمات تشکیل‌دهنده آن و برای چند جمله از حداکثر امتیاز هر جمله گرفته می‌شود. جدول (۲)، نمونه‌ای از نمره حسی اختصاص یافته‌شده به اخبار جمع‌آوری‌شده را نشان می‌دهد. این امتیازات را می‌توان به‌مثابه نماگرهایی از وضعیت واقعی اقتصاد تفسیر کرد. برای مثال، اخبار منتشرشده از تعطیلی ناشی از همه‌گیری کرونا نشانگر کاهش احتمالی ارزش افزوده بخش خدمات است. نمونه‌ای دیگر، وقتی تعداد اخبار زیادی در یک فصل درباره احتمال کاهش قیمت نفت منتشر می‌شود،

احتمالاً شاهد کاهش ارزش افزوده بخش نفت خواهیم بود. به همین ترتیب، افزایش اخبار مثبت نسبت به شاخص بورس، می‌تواند حاکی از انتظارات مثبت فعالان اقتصادی باشد.

جدول (۲): نمونه‌ای از نمره حسی اختصاص یافته به اخبار

ردیف	عنوان خبر	نمره حسی
۱	اختلاف نظر اعضای اوپک در مورد افزایش تولید نفت (۱۳۸۴/۰۱/۲۱)	-۳
۲	احتمال آغاز خاموشی محدود در برخی استان‌ها از فردا (۱۴۰۳/۰۸/۱۹)	-۴
۳	شاخص کل فقط یک پله تا ۷۱ هزار واحد (۱۳۹۴/۱۱/۶)	+۲
۴	عدم اتصال سیستمی دستگاه‌ها عامل زمان‌پر شدن اخذ مجوزها (۱۳۹۸/۱۲/۲۰)	-۱
۵	پیش‌بینی افزایش ۳۰ درصدی تولید گندم در کشور (۱۴۰۲/۰۱/۲۰)	+۳
۶	تعطیلی اصناف به علت پیک پنجم کرونا از امروز (۱۴۰۰/۰۴/۱۳)	-۵
۷	کاهش سهم بخش دولتی از حق بیمه‌ها (۱۴۰۳/۰۲/۷)	۰
۸	افزایش صادرات غیرنفتی به بیش از ۴۵ میلیارد دلار (۱۳۹۰/۰۱/۲۷)	+۲

منبع: محاسبات پژوهش

با تجمیع نمره حسی اخبار و مقیاس‌بندی با روش MIN- MAX برای هر دوره سه‌ماهه، یک سری زمانی از شاخص حسی اخبار (NSI) ایجاد شده است. قبل از تخمین مدل پژوهش ابتدا لازم است آزمون‌های نرمال بودن و مانا بودن داده‌ها به‌منزله فروض زیربنایی تخمین مدل، انجام شود. جدول (۳) نتایج این آزمون‌ها را نشان می‌دهد.

جدول (۳): نتایج آزمون‌های پیش از تخمین

نتیجه آزمون	مقدار احتمال	مقدار آماره	متغیر	فرض صفر	نوع آزمون
متغیر نرمال است.	۰/۴۷۳	۱/۵۰	Y	نرمال بودن	Jarque-Bera Test
متغیر نرمال است.	۰/۲۰۳	۳/۱۸	NSI	متغیر	
متغیر مانا است.	۰/۰۰۰	-۴/۳۸	Y	نامانا بودن	Dickey-Fuller Test
متغیر مانا است.	۰/۰۰۲	-۴/۰۳	NSI	متغیر	

منبع: محاسبات پژوهش

بعد از تأیید فرضیه‌های نرمال بودن و مانا بودن متغیرها مدل برای دوره زمانی این پژوهش، یعنی از ابتدای سال ۱۳۸۴ تا فصل سوم سال ۱۴۰۲ تخمین زده شده است. جدول

(۴) نتایج برآورد مدل را نشان می‌دهد.

جدول (۴): نتایج برآورد رشد اقتصادی با شاخص حسی اخبار ۰۳/۰۲-۱۴۰۲/۰۱-۱۳۸۴

مقدار احتمال	آماره t	انحراف معیار	ضرایب	متغیرهای مستقل
۰/۰۱۸	-۲/۴۱۶	۱/۴۱۳	-۳/۴۱۴	c
۰/۰۰۰	۹/۵۷۰	۰/۰۷۳	۰/۷۰۵	Y(-1)
۰/۰۰۴	۲/۹۲۶	۰/۲۲۶	۰/۶۶۲	NSI
۰/۶۷	R ²			

منبع: محاسبات پژوهش

با توجه به علامت و معناداری ضریب NSI، این فرضیه که داده‌های حسی را می‌توان برای کنون‌بینی رشد اقتصادی به کاربرد در سطح اطمینان ۹۹ درصد تأیید می‌شود. به عبارت دیگر می‌توان گفت برآیند احساس اخبار اقتصادی منتشرشده در رسانه‌های خبری می‌تواند در درک شرایط اقتصادی در زمان واقعی استفاده شود. زمانی که برآیند احساس اخبار منتشرشده مثبت باشد انتظار رشد اقتصادی مثبت داریم و بالعکس. به منظور اطمینان از تفسیرپذیری نتایج مدل، آزمون‌های پس از برآورد شامل نرمال بودن جملات خطا، عدم وجود خودهمبستگی و ناهمسانی واریانس در جملات خطا و عدم وجود همخطی بین متغیرهای توضیحی انجام شده است. نتایج این آزمون‌ها در جدول (۵) نمایش داده شده است.

جدول (۵): نتایج آزمون‌های پس از تخمین

نتیجه آزمون	احتمال	آماره	فرض صفر	نوع آزمون
جملات خطا نرمال است.	۰/۱۷۷	۳/۴۶۲	نرمال بودن جملات خطا	Jarque-Bera Test
خودهمبستگی بین جملات خطا وجود ندارد.	۰/۴۱۹	۰/۸۷۹	عدم خودهمبستگی بین جملات خطا	Breusch-Godfrey Serial Correlation LM Test
واریانس جملات خطا ثابت است.	۰/۶۹۲	۰/۶۰۵	عدم ناهمسانی واریانس جملات خطا	Heteroskedasticity Test: White
بین متغیرهای توضیحی هم خطی وجود ندارد.	-	۱/۱۷۴	VIF<10 عدم وجود هم خطی	VIF test

منبع: محاسبات پژوهش

برای مقایسه عملکرد مدل پیشنهادی، الگوی سری زمانی تک متغیره نیز به‌منزله مدل معیار برآورد شده است. مرتبه این الگو بر مبنای روش بکس-جنکینز به‌صورت $AR(1)$ تعیین شد. این مدل چون به مقادیر گذشته متغیر متکی است و به اطلاعات اضافه‌ای برای برآورد نیاز ندارد و همواره بعد از هر فصل مرجع (به شرط انتشار اطلاعات آمار رشد فصل قبل‌تر) قابل برآورد است، مبنای خوبی برای مقایسه سایر مدل‌های پیش‌بینی است. جدول (۶)، نتایج معیارهای ارزیابی پیش‌بینی مدل با شاخص حساسی اخبار و مدل $AR(1)$ را نشان می‌دهد.

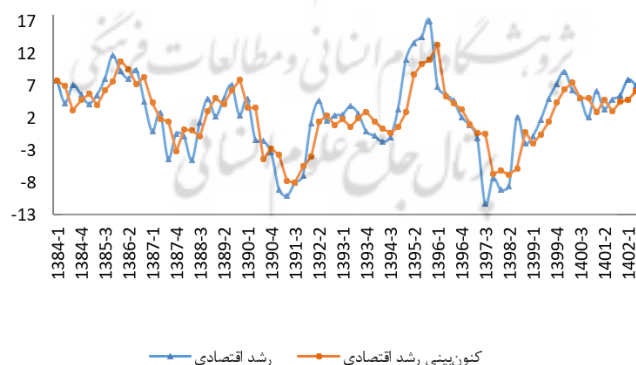
جدول (۶): معیارهای ارزیابی

نوع مدل	معیارهای ارزیابی		
	RMSE	MAE	TIC
مدل $AR(1)$ بدون شاخص حساسی	۳/۵۰	۲/۷۶	۰/۳۰
مدل با شاخص حساسی اخبار	۳/۳۰	۲/۵۶	۰/۲۸

منبع: محاسبات پژوهش

مطابق نتایج به‌دست‌آمده از جدول (۶)، استفاده از داده‌های حساسی اخبار در مقایسه با الگوی سری زمانی تک متغیره، باعث کاهش خطای پیش‌بینی می‌شود. نمودار (۱)، مقادیر واقعی رشد اقتصادی و مقادیری را که از برآورد مدل در کل دوره مورد بررسی به دست آمده است، نشان می‌دهد.

نمودار (۱): مقایسه مقادیر واقعی و مقادیر مدل کنونی بین ۱۳۸۴/۰۱-۱۴۰۲/۰۳



منبع: محاسبات پژوهش

یافته‌ها نشان می‌دهد که روندهای احساسات مثبت که از اخبار اقتصادی قابل استخراج است اغلب با دوره‌های رونق اقتصادی همسو هستند. به عبارت دیگر، افزایش تعداد اخبار اقتصادی مانند افزایش آمار اشتغال، بهبود شاخص‌های بازار سرمایه، افزایش صادرات و افزایش قیمت نفت، نشانه‌هایی از بهبود وضعیت اقتصاد و به دنبال آن افزایش رشد اقتصادی هستند. در مقابل، احساسات منفی با رکود یا عدم اطمینان همراه هستند و اخبار منفی مانند افزایش نرخ بیکاری، کاهش تولید صنعتی یا ناپایداری‌های سیاسی، به مثابه نشانه‌ای از وجود احساسات منفی در بازارها در نهایت رکود اقتصادی یا کاهش رشد اقتصادی را نشان می‌دهند. شناخت به‌هنگام این احساسات می‌تواند به کاهش نوسانات اقتصادی کمک کند. برای مثال، اگر در دوره رکود، احساسات منفی به شدت افزایش یابد، این می‌تواند نشانه‌ای از لزوم تغییر در سیاست‌های اقتصادی و مالی باشد. علاوه بر این با افزودن شاخص حسی ساخته‌شده از اخبار به مدل پیش‌بینی رشد اقتصادی فصلی، می‌توان به محض پایان هر فصل مرجع، برآوردی به‌هنگام و کارا از رشد اقتصادی همان فصل داشت. این برآورد می‌تواند به همه عوامل اقتصادی در شناسایی و مدیریت ریسک‌های مرتبط با نوسانات اقتصادی کمک کند.

۵. نتیجه‌گیری و پیشنهادها

سیاست‌گذاران، سرمایه‌گذاران، شرکت‌ها، سازمان‌ها و افراد برای تنظیم صحیح برنامه‌ها و اقدامات خود به اطلاعات به‌هنگام از شرایط اقتصاد کلان از جمله رشد اقتصادی نیاز دارند. این پژوهش نشان می‌دهد که متون خبری می‌تواند نقش مهمی در ارائه این اطلاعات داشته باشد. برخلاف آمار رسمی، اخبار به‌صورت روزانه و بدون وقفه زمانی منتشر می‌شوند و امکان دسترسی برخط به آنها در کنار توسعه بی‌وقفه روش‌های یادگیری ماشین، باعث گسترش روزافزون استفاده از این کلان‌داده ارزشمند در مطالعات اقتصادی شده است. در ایران این منبع داده به‌ندرت مورد توجه قرار گرفته است. این پژوهش در جایگاه یکی از محدود مطالعات به زبان فارسی در این حوزه از تحلیل احساس برای کمی‌سازی اطلاعات بدون ساختار متنی استفاده کرده است. در ابتدا ۳۰۱,۴۹۸ خبر اقتصادی با روش خزش از اینترنت استخراج شده است. در گام بعدی با استفاده از نسخه فارسی واژه‌نامه سنتی استرنگت، نمره یا امتیاز حسی هر خبر مشخص و با تجمیع سه‌ماهه این امتیازات، شاخص حسی فصلی ساخته شده است. نتایج نشان می‌دهد که این شاخص رابطه معنی‌داری

از لحاظ آماری با رشد اقتصادی دارد.

افزایش مجموعه داده‌های متنی که در حجم بسیار زیاد و از راه منابع اطلاعاتی مختلف در دسترس قرار گرفته‌اند در کنار توسعه روش‌های مبتنی بر یادگیری ماشین و هوش مصنوعی که مدیریت این داده‌ها را ممکن ساخته است؛ فرصت ارزیابی به‌هنگام شرایط اقتصادی را ورای داده‌های رسمی استاندارد برای همه کشورها فراهم کرده است. با این حال در کشورهای در حال توسعه که سیستم‌های آماری به اندازه کشورهای با درآمد زیاد و توسعه‌یافته قوی نیست، اهمیت داده‌های جدید و روش‌های نوین بیشتر هم می‌شود و با کمک این داده‌ها و روش‌ها می‌توان برخی شکاف‌های اطلاعاتی را کاهش داد و در نتیجه به‌طور مؤثری سیستم آماری را در ارائه آمار اقتصادی تقویت کرد. به همین علت است که تحلیل‌های متنی در اقتصاد ایران برای تخمین متغیرهایی اهمیت دارد که در زمان واقعی در دسترس نیستند یا توسط منابع رسمی اندازه‌گیری نمی‌شوند. مدل کنونی رشد اقتصادی معرفی شده در این پژوهش به بانک مرکزی و دولت کمک می‌کند تا با شناسایی به‌موقع روندهای اقتصادی، سیاست‌های پولی و مالی خود را بهینه‌سازی کنند. برای مثال، اگر پیش‌بینی‌ها نشان‌دهنده رشد اقتصادی اندک باشند، بانک مرکزی ممکن است نرخ بهره را کاهش دهد؛ همچنین این سیاست‌گذاران می‌توانند اثربخشی سیاست‌های اقتصادی گذشته و حال را بهتر ارزیابی کنند. کنونی رشد اقتصادی به‌مثابه ابزاری برای مدیریت انتظارات شرکت‌ها و سرمایه‌گذاران نیز به کاهش عدم اطمینان در مورد آینده اقتصادی کمک می‌کند که این امر به نوبه خود به کاهش ریسک و بهبود فرایند تصمیم‌گیری فعالان اقتصادی منجر می‌شود. برای پیشنهاد مطالعات بعدی می‌توان در کنار استفاده از شاخص حسی این پژوهش، کلان‌داده‌های دیگری همچون تصاویر ماهواره‌ای، داده‌های حمل‌ونقل، داده‌های مصرف انرژی و داده‌های پرداخت الکترونیکی را به منظور کاهش خطای پیش‌بینی به کار برد؛ همچنین می‌توان با استفاده از سایر رویکردهای تحلیل احساس و استفاده از مدل‌های یادگیری ماشین، عملکرد مدل این پژوهش را بهبود داد.

کنون‌بینی رشد اقتصادی فصلی ایران با استفاده از ... / بیرانوند، رزمی و ... ۱۰۳

سیاهه منابع

الف- منابع فارسی:

آهنگری آهنگرکلایی، مرتضی، علی سبطی، و مهدی یعقوبی. «ساخت واژگان به‌صورت خودکار برای تحلیل نظرات در حوزه بورس»، فصل‌نامه پردازش علائم و داده‌ها ۲۰، شماره ۲ (تابستان ۱۴۰۲):

<https://dx.doi.org/10.61186/jsdp.20.2.3>

رجبی، زینب، محمدرضا ولوی، و مریم حورعلی. «مروری بر روش‌های تحلیل احساس در متون فارسی»، فصل‌نامه پردازش علائم و داده‌ها ۱۹، شماره ۲ (تابستان ۱۴۰۱): ۱۰۷-

<https://dx.doi.org/10.52547/jsdp.19.2.107>

زارعی، عظیم، داود فیض، و غزاله طاهری. «ارائه چارچوب هوشمندی بازار اجتماعی مبتنی بر وب ۲/۰ با استفاده از تکنیک متن کاوی در وب‌سایت‌های رسانه‌های اجتماعی (مورد مطالعه:

تحلیل رقابتی در بین برندهای سامسونگ و امرسان)»، پژوهش‌های مدیریت در ایران ۲۴، شماره ۴ (۱۳۹۹): ۹۸-۱۲۵.

سوری، علی. اقتصادسنجی پیشرفته. جلد ۲، تهران: انتشارات نور علم، ۱۴۰۰.

کرامت‌فر، عبدالصمد. «مدل‌سازی چند جریانی زمینه نظرات برای تحلیل احساس» رساله دکترا، دانشگاه قم، ۱۴۰۰.

ب- منابع لاتین:

Algaba, Andres, David Ardia, Keven Bluteau, Samuel Borms, and Kris Boudt. "Econometrics meets sentiment: An overview of methodology and applications." *Journal of Economic Surveys* 34, no. 3 (2020): 512-547.

<https://dx.doi.org/10.1111/joes.12370>

Aprigliano, Valentina, Simone Emiliozzi, Gabriele Guitoli, Andrea Luciani, Juri Marcucci, and Libero Monteforte. "The power of text-based indicators in forecasting Italian economic activity." *International Journal of Forecasting* 39, no. 2 (2023): 791-808. <https://dx.doi.org/10.1016/j.ijforecast.2022.02.006>.

Ardia, David, Keven Bluteau, and Kris Boudt. "Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values." *International Journal of Forecasting* 35, no. 4 (2019): 1370-1386. <https://dx.doi.org/10.1016/j.ijforecast.2018.10.010>.

Banbura, Marta, Domenico Giannone, and Lucrezia Reichlin. "Nowcasting." in *The Oxford handbook of economic forecasting*, edited by Clements, Michael P., and David F. Hendry, 193-224. OUP USA, 2011.

- Barbaglia, Luca, Lorenzo Frattarolo, Luca Onorante, Filippo Maria Pericoli, Marco Ratto, and Luca Tiozzo Pezzoli. "Testing big data in a big crisis: Nowcasting under COVID-19." *International Journal of Forecasting* 39, no. 4 (2023): 1548-1563. <https://dx.doi.org/10.1016/j.ijforecast.2022.10.005>.
- Birjali, Marouane, Mohammed Kasri, and Abderrahim Beni-Hssane. "A comprehensive survey on sentiment analysis: Approaches, challenges and trends." *Knowledge-Based Systems* 226, 107134 (2021): 1-26. <https://dx.doi.org/10.1016/j.knosys.2021.107134>.
- Brosius, Anna, Erika J. van Elsas, and Claes H. de Vreese. "Bad news, declining trust? Effects of exposure to economic news on trust in the European Union." *International Journal of Public Opinion Research* 32, no. 2 (2020): 223-242. <https://dx.doi.org/10.1093/ijpor/edz025>.
- Cascaldi-Garcia, Danilo, Matteo Luciani, and Michele Modugno. "Lessons from Nowcasting GDP across the World." *International Finance Discussion Paper* 1385 (2023): 1-37. <http://dx.doi.org/10.17016/IFDP.2023.1385>.
- Feuerriegel, Stefan, and Julius Gordon. "News-based forecasts of macroeconomic indicators: A semantic path model for interpretable predictions." *European Journal of Operational Research* 272, no. 1 (2019): 162-175. <https://dx.doi.org/10.1016/j.ejor.2018.05.068>.
- Garz, Marcel. "Macroeconomic news: A literature survey and methodological guidelines." SSRN 4643970 (2023): 1-30 <http://dx.doi.org/10.2139/ssrn.4643970>.
- Giannone, Domenico, Lucrezia Reichlin, and David H. Small. "Nowcasting GDP and inflation: the real-time informational content of macroeconomic data releases." *ECB working paper*, no. 633 (2006): 1-52. <http://dx.doi.org/10.2139/ssrn.873658>.
- Hemmatian, Fatemeh, and Mohammad Karim Sohrabi. "A survey on classification techniques for opinion mining and sentiment analysis." *Artificial intelligence review* 52, no. 3 (2019): 1495-1545. <https://dx.doi.org/10.1007/s10462-017-9599-6>.
- Hopp, Daniel. "Benchmarking Econometric and Machine Learning Methodologies in Nowcasting." *UNCTAD Research Paper*, no. 83 (2022):1-31 <https://dx.doi.org/10.13140/RG.2.2.13344.87042>.
- Hotho, Andreas, Andreas Nürnberger, and Gerhard Paaß. "A brief survey of text mining." *Journal for Language Technology and Computational Linguistics* 20, no. 1 (2005): 19-62. <https://dx.doi.org/10.21248/jlcl.20.2005.68>.
- Keeney, Mary, Bernard Kennedy, and Joelle Liebermann. "The value of hard and soft data for short-term forecasting of GDP." *Economic Letters*, no.11 (2012): 1-6.
- Kumar, Akshi, and Mary Sebastian Teeja. "Sentiment analysis: A perspective on its past, present and future." *International Journal of Intelligent Systems and Applications* 4, no. 10 (2012): 1-14. <https://dx.doi.org/10.5815/ijisa.2012.10.01>.

- Liu, Bing. *Sentiment analysis and opinion mining*. [N.P]: Springer Nature, 2022.
- Madhoushi, Zohreh, Abdul Razak Hamdan, and Suhaila Zainudin. "Sentiment analysis techniques in recent works." *science and information conference (SAI)* (2015): 288-291. <https://dx.doi.org/10.1109/SAI.2015.7237157>.
- Manchado Marcos, Lucas. "Nowcasting with Alternative Data." Bachelor's thesis, Polytechnic University of Catalonia, 2023.
- Rambaccussing, Dooruj, and Andrzej Kwiatkowski. "Forecasting with news sentiment: Evidence with UK newspapers." *International Journal of Forecasting* 36, no. 4 (2020): 1501-1516. <https://dx.doi.org/10.1016/j.ijforecast.2020.04.002>.
- Rothman, Tiran, and Chen Yakar. "Empirical Analysis Towards the Effect of Social Media on Cryptocurrency Price and Volume." *European Scientific Journal, ESJ* 15 (2019): 31-52. <https://dx.doi.org/10.19044/esj.2019.v15n31p52>.
- Seki, Kazuhiro, Yusuke Ikuta, and Yoichi Matsubayashi. "News-based business sentiment and its properties as an economic index." *Information Processing & Management* 59, no. 2, 102795 (2022): 1-16. <https://dx.doi.org/10.1016/j.ipm.2021.102795>.
- Seo, Beomseok, Younghwan Lee, and Hyungbae Cho. "Measuring News Sentiment of Korea Using Transformer." *Korean Economic Review* 40, no. 1 (2024): 149-176.
- Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson. "Measuring news sentiment." *Journal of econometrics* 228, no. 2 (2022): 221-243. <https://dx.doi.org/10.1016/j.jeconom.2020.07.053>.
- Silverstovs, Boriss. "New York FED staff nowcasts and reality: What can we learn about the future, the present, and the past?." *Econometrics* 9, no. 1 (2021): 1-11. <https://dx.doi.org/10.3390/econometrics9010011>.
- Strycharz, Joanna, Nadine Strauss, and Damian Trilling. "The role of media coverage in explaining stock market fluctuations: Insights for strategic financial communication." *International Journal of Strategic Communication* 12, no. 1 (2018): 67-85. <https://dx.doi.org/10.1080/1553118X.2017.1378220>.
- Thelwall, Mike, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. "Sentiment strength detection in short informal text." *Journal of the American society for information science and technology* 61, no. 12 (2010): 2544-2558. <https://dx.doi.org/10.1002/asi.21416>.
- Thorsrud, Leif Anders. "Words are the new numbers: A newsy coincident index of the business cycle." *Journal of Business & Economic Statistics* 38, no. 2 (2020): 393-409. <https://dx.doi.org/10.1080/07350015.2018.1506344>.
- Tilly, Sonja, Markus Ebner, and Giacomo Livan. "Macroeconomic forecasting

through news, emotions and narrative.” *Expert Systems with Applications* 175, 114760 (2021): 1-12. <https://dx.doi.org/10.1016/j.eswa.2021.114760>.

Yue, Lin, Weitong Chen, Xue Li, Wanli Zuo, and Minghao Yin. “A survey of sentiment analysis in social media.” *Knowledge and Information Systems* 60 (2019): 617-663. <https://dx.doi.org/10.1007/s10115-018-1236-4>.

Zang, Ziqi. “Three Essays on Big Data in International Finance.” PhD diss., UCLA, 2019.

