Said Mouhtaram

GOOGLE SEARCHES, WORK AND LEISURE: PREDICTING UNEMPLOYMENT IN ITALY

ABSTRACT

This study investigates whether Google search activity can help explain and predict monthly unemployment dynamics in Italy. Using Google Trends data on job and leisure-related keywords together with official unemployment rates, we test whether online search intensity contains information about labor market conditions. For each search term, we construct up to three monthly lags to capture delayed effects of online behavior on unemployment. We then estimate several supervised machine learning models (Ordinary Least Squares, Ridge, LASSO, Random Forest, and XGBoost) using a matrix-based workflow. Regularization parameters for Ridge and LASSO are selected via rolling cross-validation, while model performance is evaluated out-of-sample on the final twelve months of data.

The results show that nonlinear models outperform purely linear ones, with Random Forest achieving the lowest out-of-sample RMSE (≈ 0.305), followed by XGBoost and Ridge regression. LASSO and OLS perform worse, reflecting overfitting and instability due to high multicollinearity among search variables. These findings indicate that the relationship between search intensity and unemployment is partly nonlinear and that moderate regularization improves predictive accuracy. Overall, the analysis extends earlier studies (Choi & Varian, 2012; D'Amuri & Marcucci, 2017) to the Italian context, highlighting the potential of online behavioral data for real-time economic monitoring.

INTRODUCTION

The growing availability of digital data has created new opportunities to monitor economic activity in real time. A seminal contribution in this area is Choi and Varian (2012), who demonstrated that Google search volumes can anticipate movements in key economic indicators such as unemployment, automobile sales, and travel patterns in the United States. Their approach, known as "predicting the present," showed that search activity can provide timely information that complements official statistics, which are typically released with delays.

Building on this idea, D'Amuri and Marcucci (2017) examined the predictive power of Google job-search queries for the U.S. labor market. Using time-series econometric models, they found that incorporating search intensity improves short-term forecasts of unemployment, particularly during periods of economic change. Their analysis, however, was limited to the United States and relied on linear forecasting methods.

This study extends that line of research to the Italian context and adopts a machine learning perspective. Using monthly ISTAT unemployment data together with Google Trends series for a wide range of job-related and leisure-related keywords, we explore whether online search behavior can help explain and predict movements in the Italian unemployment rate.

A key element of this work lies in the inclusion of leisure-related search terms alongside traditional job-related keywords. While previous research has focused almost exclusively on direct indicators of job-seeking behavior, leisure searches (such as *"cinema"*, *"discoteca"*, or *"partite di calcio"*) may also contain valuable information. The underlying intuition is that changes in employment status affect not only individuals' job search activity but also their consumption of leisure and entertainment. By comparing the predictive content of both keyword categories, the analysis provides a richer behavioral interpretation of how online activity reflects shifts in labor market conditions.

The contribution of this work is twofold. First, it applies the nowcasting framework to Italy, introducing a broader set of search terms that capture both direct job-seeking behavior and indirect lifestyle

adjustments associated with unemployment. Second, it compares the predictive performance of multiple supervised learning methods (Least Squares, Ridge, LASSO, Random Forest, and XGBoost) within a unified framework that emphasizes the bias–variance trade-off. This approach provides new evidence on the informational value of online search data and demonstrates the usefulness of modern machine learning tools for economic monitoring.

DATA DISCUSSION

Data selection

The analysis is based on monthly data for Italy drawn from two primary sources. The first source is the official unemployment rate published by ISTAT (Istituto Nazionale di Statistica), which measures the percentage of individuals in the labor force who are unemployed. These values provide a consistent, official indicator of labor market conditions over time.

The second source consists of Google Trends data, which capture the relative intensity of online searches for a set of keywords related to employment and leisure activities. Google Trends indexes search interest on a scale from 0 to 100, where 100 represents the highest recorded popularity of a term within the selected period and region. The values are relative rather than absolute, meaning they indicate how the search volume for each keyword varies with respect to its own historical maximum rather than the total number of Google searches.

The dataset includes job-related searches such as “cerco Lavoro”, “cv”, “agenzie per il Lavoro”, “INPS disoccupazione”, “cassa integrazione”, “partita iva”, “Lavoro stagionale”, “lavoro germania”, and “disoccupazione”, as well as leisure-related terms such as “cinema”, “discoteca”, “partite calcio”, “viaggio”, and “crisis economica”. All variables are recorded at monthly frequency and aligned over the same time period, so that each Google Trends observation corresponds to the unemployment rate for the same month.

The selection of search keywords was guided by both theoretical and practical considerations. Job-related queries were chosen because they directly reflect labor market behavior, capturing individuals’ job-search intensity and interest in employment opportunities or benefits. Leisure-related terms were included to test whether changes in non-work activities can indirectly signal shifts in employment status, as unemployed individuals may alter their consumption of leisure or entertainment. The number of keywords was intentionally limited to a concise and interpretable set. Expanding the list further would introduce strong collinearity, since many Google searches are highly correlated in time, and would reduce model efficiency given the limited sample size. In addition, for less common terms, Google Trends often reports incomplete or noisy series, which would undermine data quality. The chosen set therefore balances conceptual relevance, data availability, and statistical reliability, ensuring that each series contributes to the analysis without excessive redundancy.

Before the analysis, the data were cleaned and aligned to ensure consistency across all series; additionally since we are working with a time series we checked for stationarity, so we executed an ADF test on the variables:

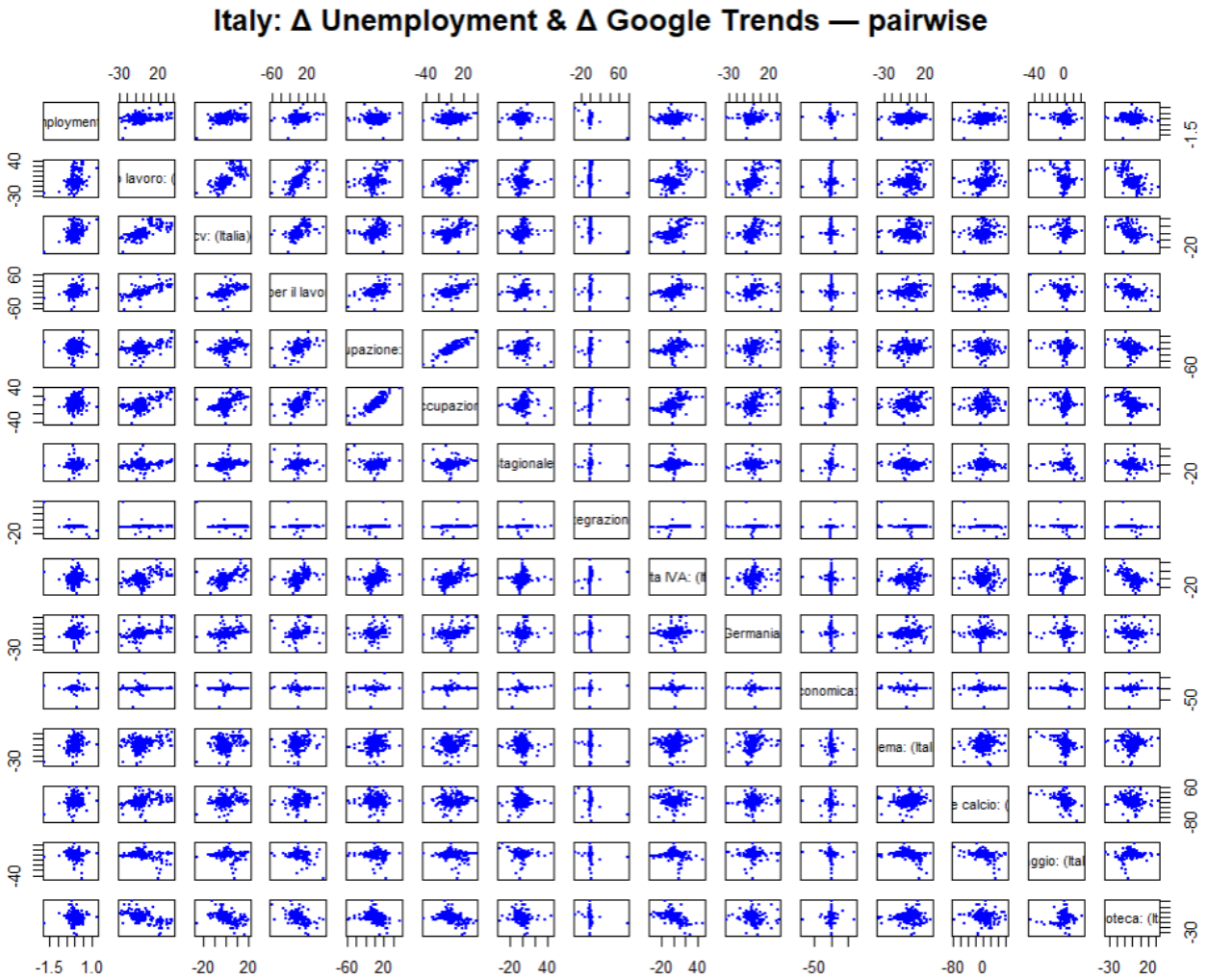
```
ADF test for Unemployment rate p-value: 0.99

ADF test for cerco lavoro: (Italia) p-value: 0.4330226

ADF test for cv: (Italia) p-value: 0.08052304

ADF test for agenzie per il lavoro: (Italia) p-value: 0.01

ADF test for disoccupazione: (Italia) p-value: 0.196726

ADF test for INPSdisoccupazione: (Italia) p-value: 0.2886794

ADF test for lavoro stagionale: (Italia) p-value: 0.01

ADF test for cassa integrazione: (Italia) p-value: 0.01

ADF test for Partita IVA: (Italia) p-value: 0.01

ADF test for lavoro Germania: (Italia) p-value: 0.6635253

ADF test for crisi economica: (Italia) p-value: 0.1456877

ADF test for cinema: (Italia) p-value: 0.01

ADF test for partite calcio: (Italia) p-value: 0.01

ADF test for Viaggio: (Italia) p-value: 0.01461803

ADF test for discoteca: (Italia) p-value: 0.6877327
```

The Augmented Dickey–Fuller (ADF) test was applied to assess the stationarity of each series in levels. The results indicate that most variables, including the unemployment rate and several Google Trends indicators (e.g. *cerco lavoro*, *cv*, *disoccupazione*), exhibit high p-values well above the 0.05 threshold, implying non-stationarity. Only a few search series (such as *agenzie per il lavoro*, *partita IVA*, *cinema*, and *cassa integrazione*) reject the null of a unit root and can be considered stationary in levels.

To ensure consistency and comparability across variables, all series were transformed into first differences (month-over-month changes). This transformation removes stochastic trends and focuses the analysis on short-term fluctuations rather than long-term levels. The dependent variable is therefore the monthly change in the unemployment rate, while the predictors are the changes in the search indices. This approach ensures that all regressors are stationary and that relationships captured by the models reflect co-movement in changes rather than common trending behavior.

Data visualization

We then plotted a scatter plot matrix to try and visualize the possible relationship between unemployment and the Google Trends variables variations.

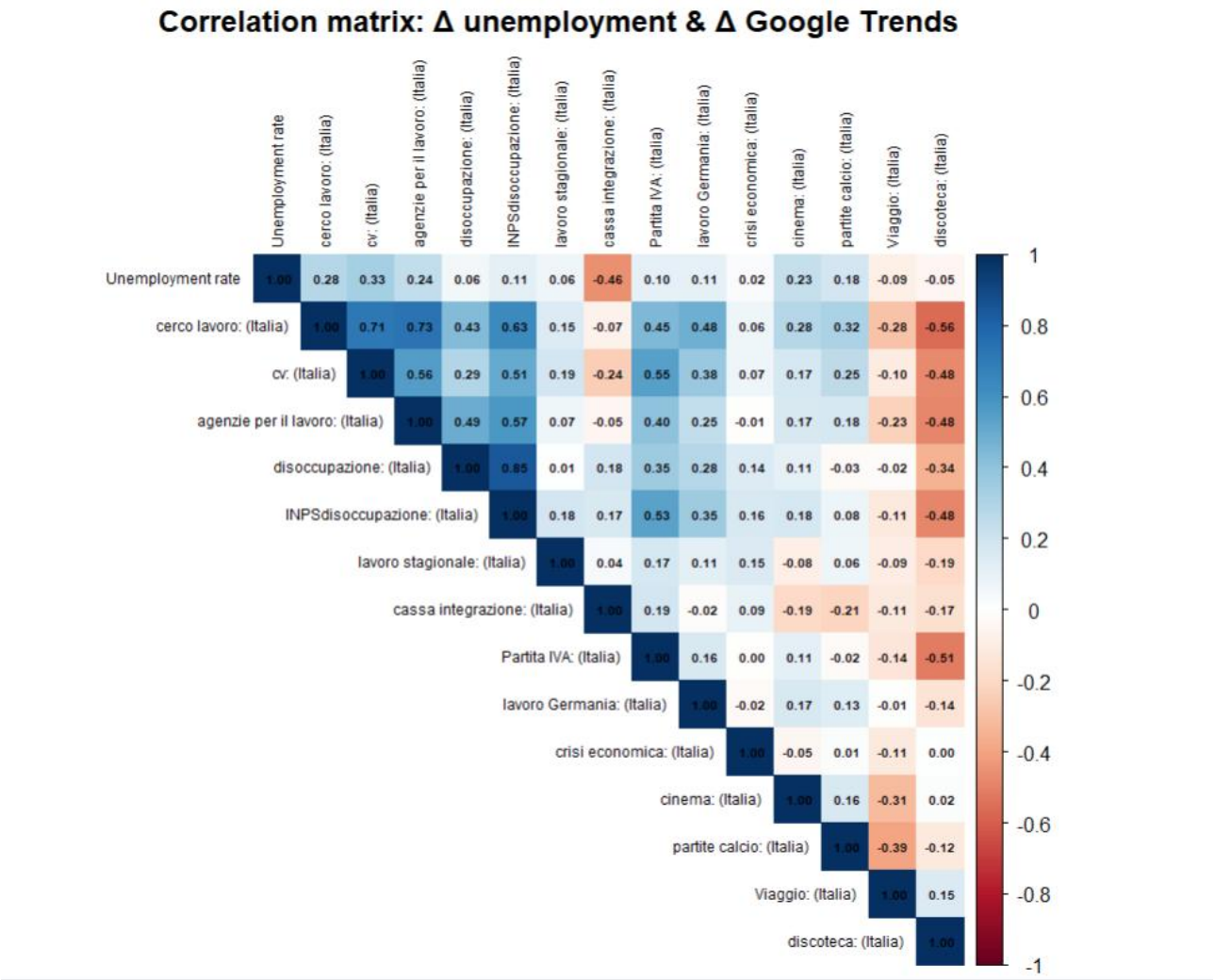


The scatterplots between monthly changes in unemployment and changes in Google search intensity show that, for most keywords, there is no clear or consistent relationship. The points are widely scattered, and only a few variables display a visible tendency toward either a positive or negative association. Some job-related searches, such as “cerco lavoro” or “disoccupazione,” show a slight positive pattern, suggesting that increases in search intensity may coincide with periods of rising unemployment. However, in most cases the relationship appears weak or dominated by noise.

For several other terms, especially those related to leisure and entertainment (such as “cinema,” “discoteca,” or “partite di calcio”) the scatterplots do not seem to indicate any systematic connection with unemployment changes. These series likely reflect short-lived attention shifts or seasonal behaviors rather than economic movements.

Overall, the visual evidence alone does not reveal strong or consistent relationships between the variables. To assess whether meaningful predictive patterns exist, it is therefore necessary to rely on formal modeling approaches. By applying both linear and non-linear methods, the analysis can determine whether any of these search variables actually contain information useful for explaining or anticipating changes in unemployment.

We also decided to visualize the correlation between the variables with a heatmap:



The correlation matrix highlights substantial co-movement among several Google search variables, particularly those directly related to job seeking. Terms such as *"cerco lavoro", "cv", "agenzie per il lavoro",* and *"disoccupazione"* exhibit high positive correlations with each other, confirming that they capture overlapping dimensions of labor market attention. In contrast, leisure-related searches such as *"cinema", "viaggio",* and *"discoteca"* are weakly correlated with both unemployment changes and job-related queries, often displaying small or negative coefficients. These patterns suggest a clear separation between employment-focused and leisure-focused behaviors but also reveal strong multicollinearity within the job-related group. This reinforces the need for regularization techniques like Ridge and LASSO, which can stabilize coefficient estimates and perform implicit variable selection in the presence of highly correlated predictors.

METHODOLOGY

This section describes each step of the analysis we made: creating lagged predictors, fitting alternative models, evaluating them through rolling cross-validation, and visualizing their performance.

Taking Lags

To capture potential delayed effects of search behavior on labor market outcomes, lagged values of the Google Trends variables are included among the predictors. This approach allows the model to detect whether changes in search intensity precede movements in unemployment by a few months. For example, increases in searches such as "cerco lavoro" or "INPS disoccupazione" might anticipate

worsening labor market conditions, while changes in leisure-related searches could reflect behavioral adjustments that occur slightly later.

In practice, three monthly lags are constructed for each Google Trends variable. This lag structure captures short-term dynamics over a quarter while maintaining a manageable number of predictors given the limited sample size. In principle, longer lag horizons (such as twelve months) could provide insight into whether search behavior exerts delayed effects on unemployment or captures persistent trends in labor market sentiment. However, extending the lag length introduces several statistical and practical challenges that outweigh the potential informational gains in this dataset.

First, each additional lag multiplies the number of predictors, since every keyword generates one new column per lag. With around fifteen Google Trends variables, a one-year (twelve-month) lag structure would result in roughly 180 predictors, in addition to the intercept term. Given that the sample contains only around 260 entries, such a design would yield a very high ratio of predictors to observations. This high dimensionality increases estimation variance and drastically reduces the reliability of OLS estimates, which become unstable or even undefined when multicollinearity is strong. Adjacent lags of the same search term are typically highly correlated (search volumes for “cerco lavoro” in March and April, for example, tend to move almost in parallel) so adding many consecutive lags contributes little new information while magnifying collinearity.

Second, the inclusion of many redundant predictors can lead to severe overfitting, particularly in flexible models such as trees or boosting algorithms. When the sample size is limited, the model may start capturing noise rather than genuine relationships, which results in good in-sample fit but poor generalization to future data. This risk is especially relevant in time-series contexts, where nearby observations are autocorrelated, reducing the effective sample size. By limiting the horizon to three lags, the analysis focuses on short-term relationships that are more likely to be robust and economically interpretable.

Third, from an economic perspective, short lags are often the most relevant for capturing behavioral responses in search activity. Individuals typically react to changes in employment conditions within a few months, not over the span of an entire year. Thus, the three-lag structure aligns with plausible behavioral dynamics while maintaining computational tractability. Regularization methods such as Ridge and LASSO could, in theory, handle higher-dimensional setups by shrinking or eliminating redundant coefficients, but doing so on a small dataset would still reduce interpretability and complicate inference.

Overall, the use of three monthly lags strikes a balance between capturing meaningful temporal dependencies and ensuring estimation efficiency. It preserves enough temporal variation to identify lead–lag patterns between search activity and unemployment while avoiding the instability, collinearity, and overfitting problems that would arise with a twelve-month lag structure. This design choice reflects a deliberate trade-off between model flexibility and statistical reliability, consistent with the bias–variance considerations that guide the broader modeling strategy.

Choosing the models

The analysis employs a range of supervised learning models to evaluate the relationship between changes in Google search activity and monthly changes in the unemployment rate. Given the uncertainty about the relationship between online behavior and labor market conditions is not well established, it is not clear whether the association is linear, stable over time, or involves complex interactions among variables. For this reason, the empirical strategy combines traditional linear models with more flexible, non-linear approaches.Linear models such as OLS, Ridge, and LASSO provide interpretable benchmarks and allow direct comparison of the magnitude and direction of predictors, while tree-based ensemble methods such as Random Forest and XGBoost offer greater flexibility to capture potential non-linearities

and interactions among variables. The use of multiple methods allows for a balanced assessment of predictive performance across different model classes.

The dataset presents several challenges that shape the choice of methods. First, the sample size is limited to fewer than 300 monthly observations, which restricts the complexity of models that can be reliably estimated. Second, many predictors are likely to be highly correlated: search queries such as *"cerco lavoro"*, *"disoccupazione"*, and *"cv"* capture overlapping job-seeking behavior and therefore move together. Third, after differencing to achieve stationarity, the relationships between variables become weaker and noisier, increasing estimation uncertainty. Finally, since the data are time ordered, model tuning must respect the temporal structure to avoid look-ahead bias. The following models are therefore selected to address these issues from different angles.

OLS

OLS serves as a benchmark model that estimates a simple linear relationship between monthly changes in unemployment and the lagged changes in Google search intensity. Because OLS includes all predictors simultaneously, it provides a useful baseline for assessing whether regularization or non-linear modeling adds predictive value. However, in this dataset, the number of predictors relative to the number of observations is relatively large, and many of the variables are correlated. These conditions make OLS estimates potentially unstable and sensitive to small changes in the data, leading to high variance and poor out-of-sample performance. The OLS model is therefore retained mainly as a reference point rather than as a reliable predictive tool.

Ridge

Ridge regression introduces an L_2 penalty on the size of the coefficients, shrinking them toward zero and thereby reducing the sensitivity of the model to multicollinearity. This feature is particularly relevant here, as many search terms convey similar information about labor market conditions. By penalizing large coefficients, Ridge balances bias and variance, allowing the model to retain all predictors but with dampened effects. The penalty parameter λ is selected through a rolling-origin cross-validation procedure that evaluates one-step-ahead forecast errors over expanding windows, the same thing was done for LASSO. This ensures that parameter tuning respects the time ordering of the data and avoids using future information to predict the past. This procedure effectively embeds model tuning within a forecasting framework, where each λ value is chosen based solely on information available up to that point in time. As a result, the evaluation mimics how the model would perform in a genuine real-time setting, producing a more reliable estimate of out-of-sample predictive accuracy. Moreover, by repeating the process across multiple rolling windows, we obtain a distribution of optimal λ values rather than a single estimate, allowing us to identify a penalty level that performs robustly over different historical periods. This conservative approach limits the risk of overfitting to any particular subsample and ensures that the selected amount of regularization reflects stable predictive structure rather than transitory noise in the data.

LASSO

LASSO regression extends the idea of penalization by applying an L_1 norm, which can shrink some coefficients exactly to zero. This property makes LASSO especially suitable for high-dimensional and correlated predictors, as it performs automatic variable selection. In this context, it allows the model to identify which search queries (and lags) carry the most predictive information about short-term movements in unemployment, while discarding those that add noise. This is particularly valuable given the limited sample size and the risk of overfitting. As with Ridge, the regularization strength is tuned using rolling cross-validation to ensure temporal consistency in model selection.

Random Forest

The Random Forest model provides a non-linear alternative capable of capturing complex relationships and interactions between variables. Each tree in the forest is trained on a bootstrap sample, and a random subset of predictors is considered at each split. This randomization helps reduce overfitting and handle correlated features (two key issues in this dataset). Random Forests do not assume any specific functional form, making them robust to non-linear or threshold effects that might arise in behavioral data such as Google Trends. However, because they rely on averaging many trees, they may smooth out short-term variations and underperform when relationships are weak or highly volatile, as may occur in the differenced data.

In implementation, the Random Forest was fit using the randomForest package with ntree = 500 and mtry = √p (where p is the number of predictors) during each iteration of the rolling-origin evaluation. Trees were grown unpruned with the package defaults for regression (bootstrap sampling with replacement and nodesize = 5), while all other options were left at their defaults. Given the relatively small sample and the high computational cost of repeated refitting, we adopted these well-established default settings, which are widely recommended for regression tasks and provide a good balance between accuracy and efficiency. The choice of $mtry = \sqrt{p}$ promotes tree diversity in the presence of correlated predictors, and using 500 trees ensures prediction stability without excessive runtime. This configuration maintains temporal consistency across rolling iterations and keeps the procedure computationally tractable, while still enabling the model to capture meaningful non-linear relationships among lagged search indicators.

XGBoost

XGBoost represents a more adaptive non-linear approach that builds trees sequentially, where each new tree corrects the errors made by the previous ones. This boosting mechanism allows the model to focus on difficult-to-predict observations and capture subtle patterns that linear models cannot detect. Given that relationships in the differenced data may be weak or irregular, boosting offers a way to extract small predictive signals by combining many simple models. To mitigate overfitting given the small dataset, conservative parameter choices are used (300 rounds, learning rate 0.1, moderate tree depth, and subsampling). For tree-based methods (Random Forest and XGBoost), standard hyperparameter values were adopted rather than re-tuned within each rolling window. Given the limited sample size and the high computational cost of nested rolling validation, parameters such as the number of trees, tree depth, and learning rate were set to commonly used defaults (Breiman, 2001; Chen & Guestrin, 2016). This choice ensures stable estimation while preventing overfitting to small validation samples.

Model Evaluation

Model evaluation follows a rolling-origin (expanding-window) cross-validation framework, which mimics real-time forecasting conditions. At each step, the model is trained on all available observations up to time t and used to predict the following month (t + 1). Prediction errors from these one-step-ahead forecasts are then averaged over time to obtain a rolling RMSE, providing a time-consistent measure of out-of-sample performance for each model class (OLS, Ridge, LASSO, Random Forest, and XGBoost).

We must highlight that while the rolling-origin framework offers a realistic and statistically sound way to evaluate predictive accuracy, it is also computationally demanding, especially in settings involving multiple algorithms and hyperparameter tuning. Because each step of the rolling procedure requires refitting the model on an expanding sample, the total number of model estimations grows roughly linearly with the number of time periods. For simple linear models such as OLS or Ridge, this additional cost is minor; however, for non-linear ensemble methods like Random Forests or XGBoost (where hundreds of trees are trained at each iteration) the computational burden can become substantial. This creates a clear trade-off between methodological rigor and computational feasibility. In principle, lighter

evaluation schemes (e.g. a single train–test split or k-fold cross-validation) could reduce runtime, but they would violate the temporal ordering of the data and risk look-ahead bias, yielding overly optimistic performance estimates. The choice to retain a full rolling-origin setup thus reflects a deliberate preference for forecasting validity over computational convenience: even though it requires significant processing time, it ensures that every model is evaluated under conditions that closely replicate real-time nowcasting, where only past information is available when predicting the next observation.

Visualization

To visualize model performance in a way that is easy to interpret, we also run a prediction on the most recent 12 periods. This is not meant to replace the full rolling-origin evaluation (which already simulates true nowcasting by expanding the training sample and forecasting t+1 at every step); instead, it gives a clear side-by-side visual comparison of how the different models track the most recent dynamics of Italian unemployment.

FINDINGS

	Model	Rolling_RMSE
1	OLS	0.3757908
2	Ridge	0.3487471
3	LASSO	0.3740392
4	Random Forest	0.3047529
5	XGBoost	0.3448284

The comparison of out-of-sample performance based on the rolling-origin evaluation shows that the nonlinear Random Forest model delivers the lowest prediction error (RMSE ≈ 0.305), outperforming all other approaches. Gradient boosting (RMSE ≈ 0.345) and Ridge regression (RMSE ≈ 0.348) form the next tier, while OLS (RMSE ≈ 0.376) and LASSO (RMSE ≈ 0.374) are less accurate. This ranking suggests that, although regularization helps compared to plain OLS in many macro–nowcasting settings, in this dataset the biggest gains actually come from allowing nonlinear, interaction-driven structure rather than purely shrinking linear coefficients.

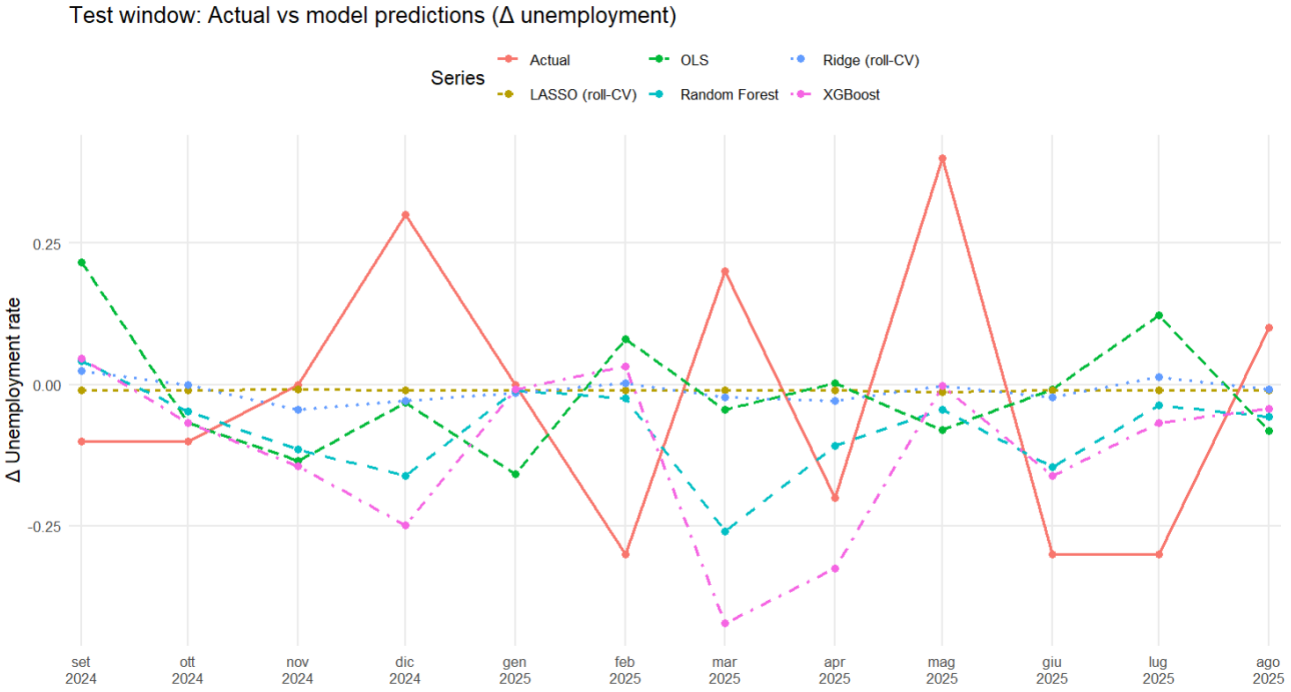
The Random Forest is the best performing because the feature set is high-dimensional (multiple Google Trends terms with multiple lags) and potentially noisy, but also plausibly nonlinear: search intensity for certain keywords may only matter after unemployment has already started moving, or only in combination with certain lagged terms of unemployment itself. Tree ensembles are good at capturing that kind of interaction and threshold behavior without having to specify it manually. At the same time, the forest's built-in averaging across many trees keeps variance under control, which prevents catastrophic overfitting even in a relatively short macro time series.

Ridge regression's performance being competitive with XGBoost and clearly better than OLS is still informative. OLS does no shrinkage, so with many correlated predictors (the correlation matrix shows that a lot of the Google search series move together), coefficient estimates become unstable and chase noise, which hurts generalization. Ridge (L_2 penalty) pulls coefficients toward zero in a smooth way, which stabilizes them under multicollinearity and lowers variance. That's why Ridge beats OLS.

LASSO (L_1 penalty) also regularizes, but it enforces sparsity, it tries to select a small subset of predictors and zero out the rest. In very short samples, that hard selection step can actually throw away weak but real signals that Random Forest is still able to exploit in combination, and that Ridge can partially retain via soft shrinkage. That helps explain why here LASSO does not outperform Ridge.

Overall, the evidence here is that unemployment dynamics in differences are not purely linear in lagged predictors: allowing nonlinear splits and interactions (Random Forest) gives the best nowcasting accuracy, while pure OLS is too unstable and aggressively sparse selection (LASSO) is slightly too aggressive for this sample size.

Prediction of the last 12 months



The figure compares the performance of the five models (OLS, Ridge, LASSO, Random Forest, and XGBoost) over the 12-month test window, where the dependent variable represents monthly changes in the unemployment rate (Δ unemployment). The actual series exhibits clear short-term fluctuations, alternating between positive and negative variations, while most models display smoother dynamics centered around zero. This indicates the difficulty of capturing short-term changes in unemployment once long-term trends have been removed through differencing.

Among the linear specifications, Ridge regression and LASSO deliver nearly flat predictions, indicating strong shrinkage that stabilizes the estimates but also underreacts to short-term fluctuations. OLS, while slightly more responsive, still fails to capture the direction and magnitude of the largest swings (e.g., December 2024 and May 2025).

The Random Forest and XGBoost models show greater variation and occasionally move in the correct direction of the true series, suggesting some ability to detect nonlinear shifts; however, they tend to overshoot or lag behind peaks and troughs. Overall, no model perfectly tracks the pronounced spikes in unemployment changes, highlighting the challenge of short-horizon nowcasting in such a volatile, noise-dominated time series.

CONCLUSIONS

This study explored whether Google search activity can improve the short-term prediction of unemployment dynamics in Italy. Building on the "predicting the present" framework introduced by Choi and Varian (2012) and extended by D'Amuri and Marcucci (2017) for the United States, we examined both job-related and leisure-related search terms to assess whether online behavior contains timely information about labor market conditions. Unlike earlier research focused on linear econometric models, our analysis adopted a machine-learning approach that compares linear, regularized, and nonlinear predictive models within a unified rolling-origin framework.

Our empirical findings align with and extend previous evidence in several ways. Similar to prior studies, we confirm that search-based indicators can contain limited but measurable information about short-term unemployment fluctuations. However, in contrast with F. D'Amuri and J. Marcucci's results (where linear models with Google queries significantly improved forecasts) we find that in the Italian case,

nonlinear methods such as Random Forest perform best. This suggests that the relationship between search intensity and unemployment is not purely linear and may involve threshold effects or interactions between search categories and lagged unemployment changes. The strong performance of Random Forest, followed by XGBoost and Ridge regression, highlights the value of allowing flexible structures that can accommodate complex dependencies in behavioral data.

At the same time, the weak predictive power of LASSO and OLS emphasizes the limits of linear sparsity-based approaches when the signal is faint and the predictors are highly correlated. The correlation matrix revealed that many job-related searches move together, leading OLS to overfit noise and LASSO to discard potentially useful information. Ridge's smooth shrinkage achieved better bias–variance balance, confirming that regularization remains valuable even when nonlinear methods outperform purely linear models.

From a substantive perspective, our analysis contributes to the literature by applying this framework to Italy, incorporating both job-related and leisure-related search terms that help capture aspects of household behavior during labor market adjustments. This broader perspective helps link search data to consumer sentiment and lifestyle changes beyond the job search itself.

Overall, the results suggest that Google search data provide some incremental but limited predictive gains for monthly unemployment changes once underlying trends are removed. The modest reduction in forecast error highlights both the promise and the limitations of digital trace indicators for macroeconomic nowcasting in small samples. In a low-signal, noise-dominated setting like monthly Italian unemployment, regularization and nonlinear learning methods help stabilize predictions and uncover weak relationships that standard regressions miss.

Future research could expand this analysis by incorporating additional categories of online activity (such as mobility or social-media indicators), testing higher-frequency data (weekly Google Trends), and integrating mixed-frequency or Bayesian models to exploit temporal structure more efficiently. As digital data sources continue to proliferate, combining them with traditional economic indicators through machine learning will remain a promising avenue for improving the timeliness and robustness of economic monitoring.

REFERENCES

Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of KDD 2016, 785–794.

Choi, H., & Varian, H. (2012). Predicting the Present with Google Trends.

D'Amuri, F., & Marcucci, J. (2017). The predictive power of Google searches in forecasting unemployment. International Journal of Forecasting.

Askitas, N., & Zimmermann, K. F. (2009). *Google econometrics and unemployment forecasting. Applied Economics Quarterly, 55*(2), 107–120.

Varian, H. R. (2014). *Big data: New tricks for econometrics. Journal of Economic Perspectives, 28*(2), 3–28.

DATA SOURCES

ISTAT

Google Trends