

Systemic Analysis, Complexity, and Sensitivity of the GoDaddy Microbusiness Density Forecasting System

Daniel Felipe Gómez Miranda

Dept. of Computer Engineering

Universidad Distrital Francisco José de Caldas

Email: {dfgomezm}@udistrital.edu.co

Julian David Cabrera Barragan

Dept. of Computer Engineering

Universidad Distrital Francisco José de Caldas

Email: {jdcabrerab}@udistrital.edu.co

Andrés Julián Vargas Medina

Dept. of Computer Engineering

Universidad Distrital Francisco José de Caldas

Email: {ajvargasm}@udistrital.edu.co

Geraldine Alejandra Vargas Moreno

Dept. of Computer Engineering

Universidad Distrital Francisco José de Caldas

Email: {geavargasm}@udistrital.edu.co

Abstract—This paper analyzes the GoDaddy Microbusiness Density Forecasting competition through a systemic perspective. The study identifies the system’s structure, sources of complexity, sensitivity factors, and the role of chaos and randomness. The analysis highlights strengths such as dataset relevance and evaluation metrics, while noting weaknesses related to noise, missing variables, and chaotic behavior inherent to socio-economic and natural systems.

Index Terms—Systemic analysis, complexity, sensitivity, chaos, forecasting, microbusiness density

I. COMPETITION SUMMARY

The *GoDaddy Microbusiness Density Forecasting* competition aims to predict microbusiness density at the local level. The dataset includes time series by region with demographic, economic, and social variables, along with historical records of microbusiness density.

Key constraints include:

- Incomplete and noisy historical data.
- Limited coverage: relevant predictors are missing.
- Uncertainty due to uncontrollable external shocks such as crises or pandemics.

II. SYSTEM ANALYSIS REPORT

A. Systemic Analysis

The system can be understood as a socio-economic ecosystem with multiple interdependent factors:

- **Purpose:** predict microbusiness density to support policy, investment, and planning.
- **Components:**
 - a) Demographic and social variables (population, migration, education).
 - b) Economic variables (income, unemployment, credit, infrastructure).
 - c) Institutional environment (regulations, taxes, government support programs).
 - d) External factors (economic crises, pandemics, technological shocks).
 - e) Predictive models based on historical data.

- **Inputs:** socio-economic and historical datasets.
- **Processes:** statistical or machine learning models.
- **Outputs:** microbusiness density forecasts by region and period.
- **Feedback:** comparison of predictions with observed values for model retraining and correction.

B. Diagrams

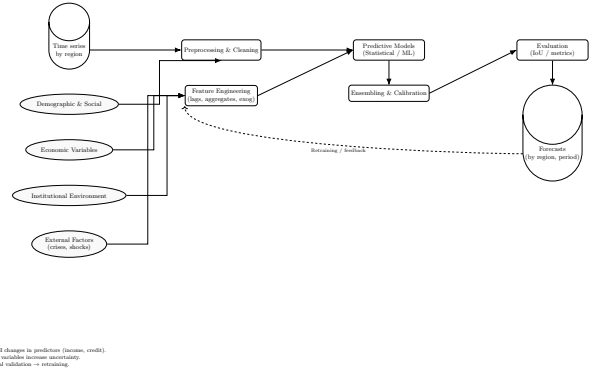


Fig. 1. Systemic workflow for microbusiness density forecasting: from raw time series and socio-economic variables to forecasts and retraining.

C. Complexity and Sensitivity

The system exhibits the following characteristics:

- **Complexity:** multicausality, nonlinearity, cumulative effects, high dimensionality, and uncertainty in the data.
- **Sensitivity:**
 - a) Small changes in key predictors (income, credit, policies) can strongly alter outcomes.
 - b) Hyperparameter tuning significantly affects accuracy and the risk of overfitting.
 - c) Input errors propagate and amplify over time in forecasts.
 - d) Thresholds exist where minor variations cause disproportionate effects.

D. Chaos and Randomness

A parallel example is the *Google Research – Identify Contrails* competition, linked to atmospheric systems:

- **Atmospheric chaos:** small variations in humidity or wind lead to very different outcomes.
- **Randomness in data:** sensor noise and labeling errors add uncertainty.
- **Nonlinear feedback:** contrails may grow into larger cloud systems.
- **Butterfly effect:** flight paths or engine types introduce unpredictable variability.

This illustrates that chaos and randomness cannot be eliminated but must be managed through robust modeling.

III. CONCLUSION

The analysis highlights both strengths and weaknesses:

- **Strengths:** problem relevance, extensive historical data, and a clear evaluation metric (IoU).
- **Weaknesses:** data noise, missing variables, sensitivity to external conditions, and limits imposed by chaotic system behavior.

In summary, predicting microbusiness density requires models capable of handling uncertainty, adapting to perturbations, and generalizing beyond the training dataset.

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.