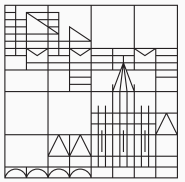
Universität Konstanz



Deep Learning for the Social Sciences

Education Investment Forecasting for Regional Economic Development

July 16, 2025

Outline

1. Motivation
2. Data
3. Methodology
4. Outlook

Motivation

Motivation

- Education is a major public investment: US Federal, state, and local governments spend over **$857 billion** annually on K–12 education, around **5.59%** of GDP.
- This funding supports approximately **49.6 million K–12 students** and **13.5 million postsecondary students** across the United States.
- There are significant disparities in **per-student spending** across states and counties, both in total amounts and in allocation categories (see next slide).

Motivation

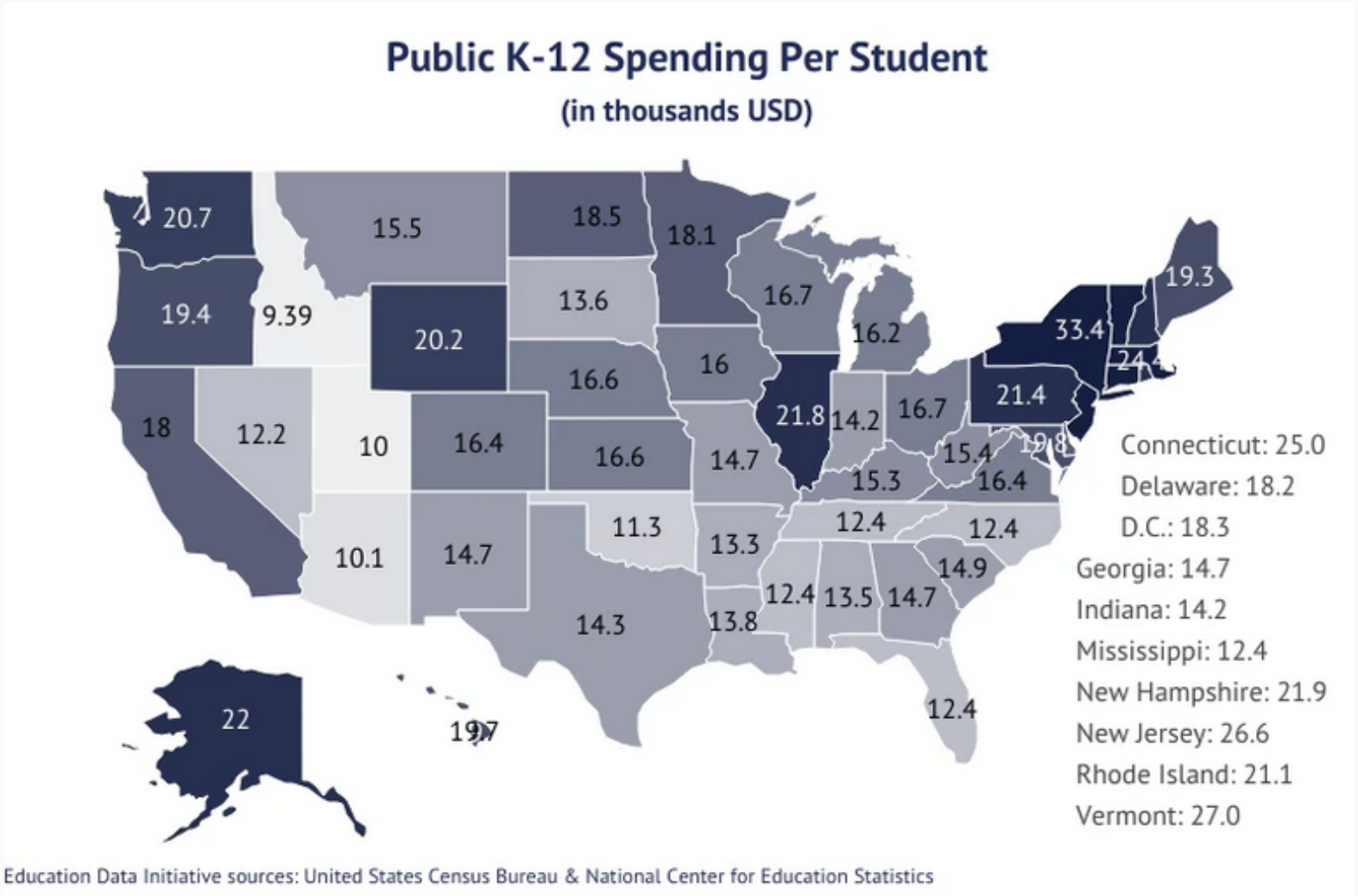


Figure 1: Per-student education spending varies significantly across regions.

Source: educationdata.org

Motivation

- Using an **RNN model** to forecast economic outcomes 10 years ahead based on education spending can provide valuable insights into **how and where to allocate education funding** most effectively.
- This approach may help identify strategies to **bolster regions with lower economic development** through targeted education investments.

Data

Data Collection

Metropolitan Areas

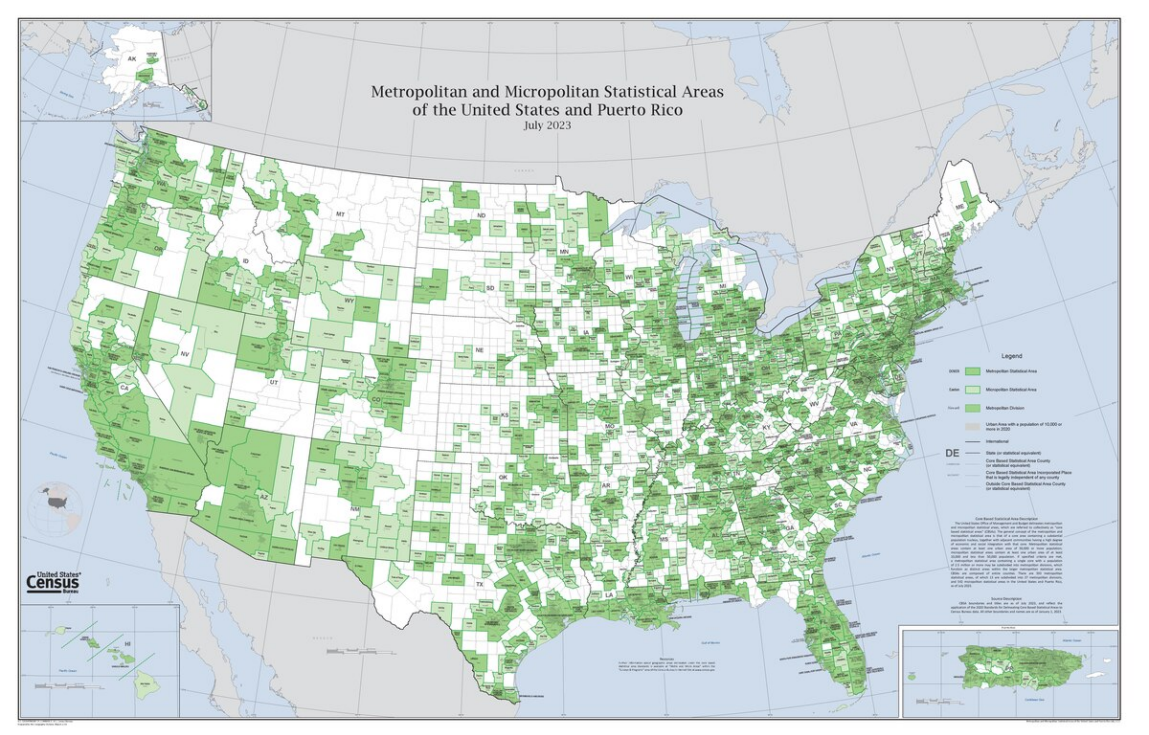


Figure 3: Metropolitan areas of the United States

- As of 2024, **282 million** U.S. citizens live in metropolitan areas, constituting **83%** of the total population.

Source: wikimedia.org

Data Collection

(Teacher) Salary Data

- Data source is the US Bureau of Labor Statistics (BLS) for metropolitan area level data from 1997 to 2024.
- The database includes occupational employment and wage statistics like mean and median hourly wages next to employee numbers (per profession categorized by standardized OCC-codes), which will further be filtered to only education-related occupations and professions.
- Hereby, the inclusion of other education related professions is also possible.

Data Collection

K–12 Education Expenditure Data

- Data source is the National Center for Education Statistics (NCES), which provides district-level education data from 1995 onward. District data will be aggregated to the county level and further grouped into the metropolitan areas.

- The database includes information on enrollment by age, gender, and ethnicity; the number of teachers and staff; general finance data; as well as revenues and expenditures.

- Expenditures are categorized by function (e.g., instruction, administration, maintenance), salary data (redundant), and other categories such as adult education, construction, and equipment.

Data Collection

Economic Outcomes Data

- **Per capita personal income** from the U.S. Bureau of Economic Analysis (counties or metropolitan areas, 1969-2023)
- **Annual average of monthly employment levels** from the U.S. Bureau of Labor Statistics (counties or metropolitan areas, 1975-2024)
- **Average weekly wages** from the U.S. Bureau of Labor Statistics (counties or metropolitan areas, 1975-2024)

Data Collection

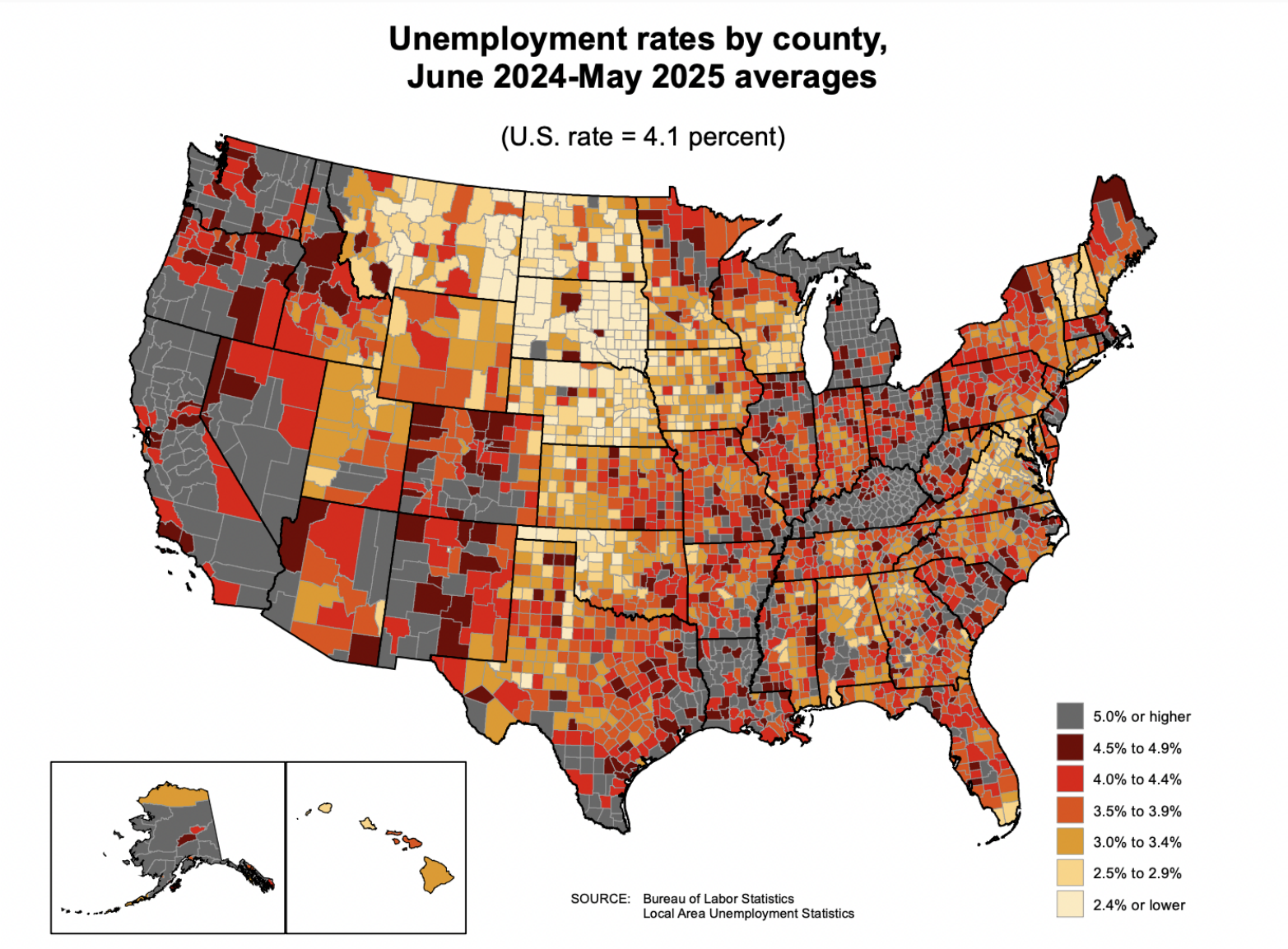


Figure 2: Unemployment rates vary significantly across counties.

Methodology

LSTM Model

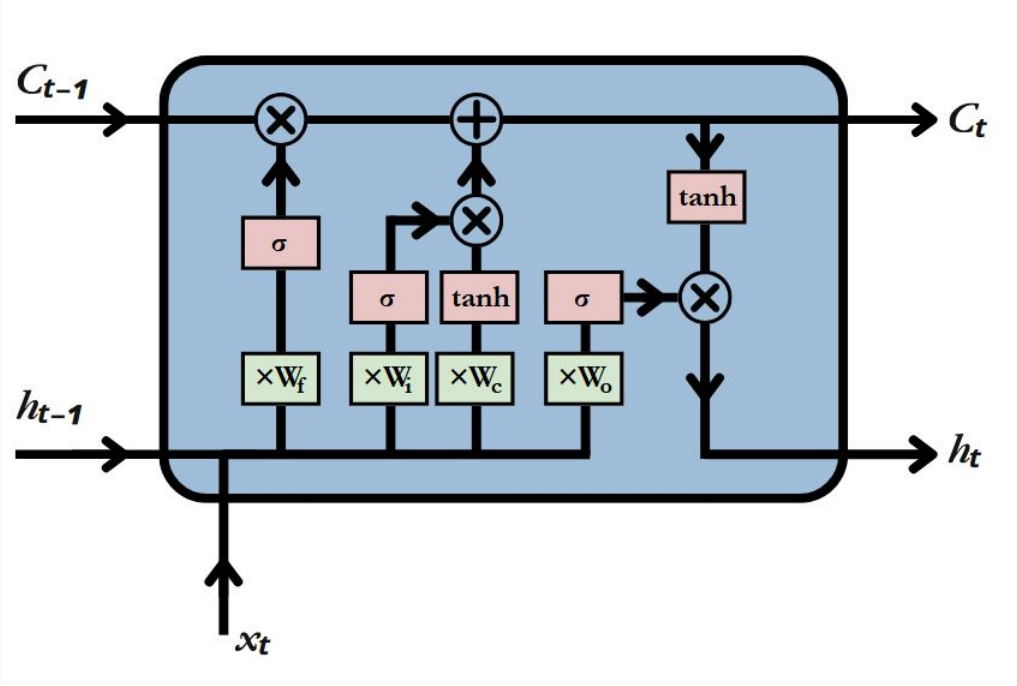


Figure 3: Long Short-Term Memory (LSTM) cell architecture

- Multivariate LSTM model trained on education expenditure indicators and some regional economic indicators
- Designed to forecast long-term outcomes with a time horizon of 10 years

GRU Model

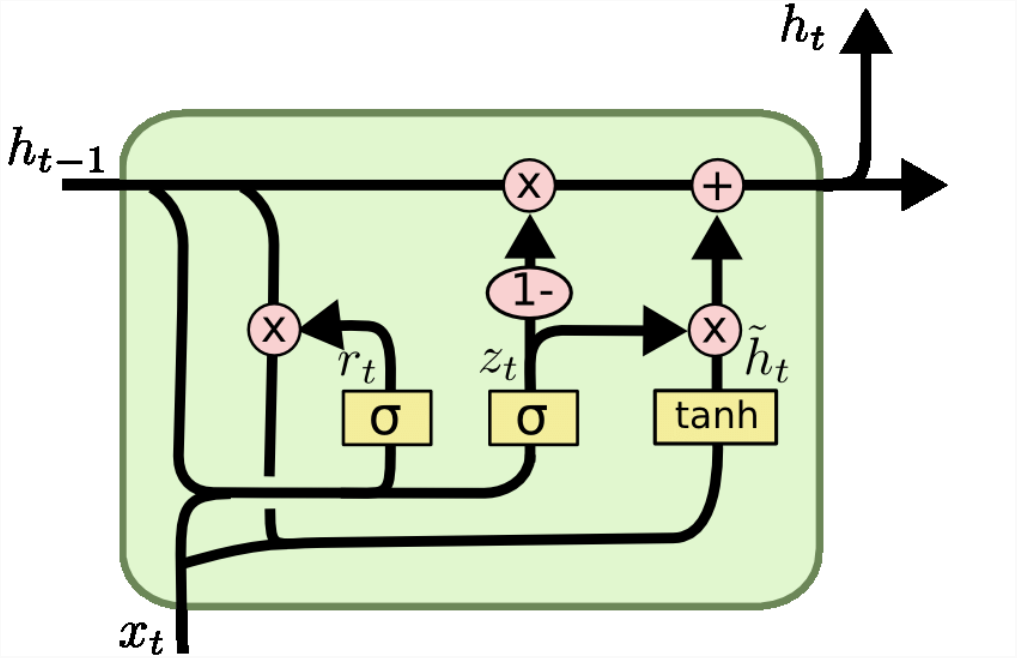


Figure 4: Gated Recurrent Unit (GRU) cell architecture

- Simplified version of the more complex LSTM
- Fewer parameters compared to LSTMs, leading to faster training and potentially better performance on some tasks

Outlook

Outlook

In Sum

- Finalize data preprocessing pipelines
- Align multi-source time series across a 25-30 year historical window
- Train and validate best performing model
- Identify which categories of education spending are most predictive of future regional economic indicators
- Provide policy-relevant insights that can help guide strategic investments in education

Outlook

Possible Challenges

- Limited availability of county-level education expenditure data, potentially restricting the analysis window to 1997 onwards.
- This limitation results in a relatively short observation window, which may constrain the model's ability to forecast 10–15 years into the future.

References

- Hanson, Melanie. *U.S. Public Education Spending Statistics*. EducationData.org, February 8, 2025. https://educationdata.org/public-education-spending-statistics
- U.S. Bureau of Labor Statistics, Local Area Unemployment Statistics, "Unemployment rates by county, latest 12-month averages". https://www.bls.gov/lau/tables.htm
- U.S. Bureau of Labor Statistics, Occupational Employment and Wage Statistics, "Occupational Employment and Wage Statistics Tables." https://www.bls.gov/oes/tables.htm
- U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency (School District) Universe Survey". https://nces.ed.gov/ccd/elsi/tableGenerator.aspx