TRADITIONAL STATISTICAL METHODS VS. NEURAL NETWORKS

Case Worker Employment Mobility in Child Protection Services Regions

ANDREW S. QUINN, LMSW

UNIVERSITY OF TEXAS AT ARLINGTON
GOALS

- EXPLAIN PURPOSE OF PRESENTATION
- EXPLAIN A NEURAL NETWORK
- EXPLAIN CPS DATA
- EXPLAIN LOGISTIC REGRESSION ANALYSIS
- COMPARE METHODS
- CONCLUSION
NEURAL NETWORKS: AN INTRODUCTION
BRAIN
THREE LAYERED NEURAL NETWORK

Data is passed through the network by the ways of statistical functions

Input Layer  Hidden Layer  Output

4-3-1
THE OVERALL PURPOSE OF THE NETWORK IS TO CAPTURE THE IMPORTANT FEATURES AMONG INPUT VARIABLES AND THEN MAP THEM TO A SET OF OUTPUT VARIABLES
HOW IS THIS DONE

REDUCING THE ERROR BETWEEN OBSERVED OUTCOMES AND PREDICTED OUTCOMES
BACK PROPAGATION

REDUCING ERROR

Input Layer  Hidden Layer  Output
THE CASE FOR NEURAL NETWORKS

- **MASSIVE PARALLELISM**: may utilize hundreds of thousands of neurons
- **HIGH INTERCONNECTIVITY**: multiple connections between neurons
- **COLLECTIVE COMPUTATION**: problems are solved by joint activities of all the neurons in the system
- **SELF-ORGANIZATION**: capable of changing structure to reflect new input patterns
BENEFITS

- HANDLES NONLINEARITY IN THE DATA
- NO DISTRIBUTIONAL ASSUMPTIONS
- ABILITY TO HANDLE MISSING AND NOISEY DATA
- MULTICOLLINEARITY
- CAN LEARN TO RECOGNIZE PATTERNS
<table>
<thead>
<tr>
<th>PSYCHOLOGY</th>
<th>POLITICAL SCIENCE</th>
<th>SOCIOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPTH PERCEPTION</td>
<td>MUNICIPAL JURI SPRUDENCE</td>
<td>PREDICTION OF CHILD ABUSE</td>
</tr>
<tr>
<td>ENHANCING EXPERT SYSTEMS</td>
<td>PREDICTION OF THE LIKELIHOOD OF PASSING THE BAR</td>
<td>PREDICTING WHITE COLLAR CRIME</td>
</tr>
<tr>
<td>DEVELOPMENT OF ARTIFICIAL INTELLIGENCE</td>
<td>MODELING DATA ON THE HISTORY OF MILITARY COMBAT</td>
<td>PREDICTING HUMAN DECISION MAKING IN RELATION TO SOCIOLOGICAL THEORIES OF RELIGION</td>
</tr>
<tr>
<td>MODELING MEMORY</td>
<td>ANALYZE VOTING BEHAVIOR</td>
<td>HUMAN SERVICE APPLICATIONS</td>
</tr>
<tr>
<td>SPOKEN WORD RECOGNITION</td>
<td>IDENTIFYING LEGAL PRECEDENTS FOR CASES</td>
<td>PREDICTING VIOLENT CRIMINAL BEHAVIOR</td>
</tr>
</tbody>
</table>
TURNOVER
TURNOVER

VOLUNTARY OR INVOLUNTARY TERMINATION OF EMPLOYMENT
RESEARCH QUESTION

WHAT ARE THE EMPLOYMENT PATHS AND CHARACTERISTICS OF CASE WORKER / SUPERVISOR TURNOVER AND RETENTION
OUR DATA

N=536
OUTPUT VARIABLE

LEAVE OR STAY
PREDICTOR VARIABLES

- AGE
- DEGREE
- EDUCATION
- ETHNICITY
- GENDER
- CURRENT CLASSIFICATION (WORKER OR SUPERVISOR)
- EXPERIENCE
- PRIOR EXPERIENCE
- CURRENT COUNTY OF EMPLOYMENT
- CURRENT JOB TYPE (INTAKE/INVESTIGATION OR OTHER)
- TOTAL ASSIGNMENTS POST TRAINING
- EMPLOYMENT COUNTY OF FIRST JOB ASSIGNMENT
- JOB TYPE OF FIRST JOB ASSIGNMENT
- LENGTH OF TENURE AT FIRST JOB ASSIGNMENT
- EMPLOYMENT COUNTY OF SECOND JOB ASSIGNMENT
- JOB TYPE OF SECOND JOB ASSIGNMENT
- LENGTH OF TENURE AT SECOND JOB ASSIGNMENT
UTILIZING THE MODELS
LOGISTIC REGRESSION

A MULTIVARIATE TECHNIQUE FOR ESTIMATING THE PROBABILITY THAT GIVEN THE PRESENCE OR ABSENCE OF AN INDEPENDENT VARIABLE, THE DEPENDENT VARIABLE WILL PREDICT IN A CERTAIN DIRECTION
BAYESIAN NETWORK

EXAMINES HOW PROBABILITIES ARE AFFECTED BY WHAT WE KNOW ABOUT A SITUATION BEFORE AND AFTER WE EXAMINE THE DATA
DATA ALLOCATION

- **Training Dataset**: The data set used to build the model
- **Validation Dataset**: Prevents the network from over training
- **Test Dataset**: Measures the performance of a trained network and its ability to generalize
BENEFITS TO USING A BAYESIAN NETWORK

USES AN ERROR TERM INSTEAD OF A VALIDATION SET ALLOWING FOR MORE DATA TO BE ALLOCATED FOR THE TRAINING DATA SET
DATA ALLOCATION

<table>
<thead>
<tr>
<th>Training</th>
<th>Validation</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>70%</td>
<td>N = 375</td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td>N = 161</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Percentage</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Stayers</td>
<td>94%</td>
<td></td>
</tr>
<tr>
<td>Leavers</td>
<td>83%</td>
<td></td>
</tr>
</tbody>
</table>

**NEURAL NETWORKS**
# Logistic Regression*

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Stayers</th>
<th>Leavers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>76%</td>
<td>86%</td>
<td>54%</td>
</tr>
</tbody>
</table>

The same 70% of data allocated for a Bayesian network’s training data.
## Comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>Overall</th>
<th>Stayers</th>
<th>Leavers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>76%</td>
<td>86%</td>
<td>54%</td>
</tr>
<tr>
<td>Bayesian Network</td>
<td>90%</td>
<td>94%</td>
<td>83%</td>
</tr>
</tbody>
</table>
ARE NEURAL NETWORKS STATISTICAL TOOLS?

• TRANSFER FUNCTIONS THAT NEURAL NETWORKS USE ARE STATISTICAL

• THE PROCESS OF ADJUSTING WEIGHTS (passing data through the network) TO ACHIEVE A BETTER FIT TO THE DATA USING WELL-DEFINED ALGORITHMS Follows FROM STATISTICAL THEORY
SOME FINAL THOUGHTS

• WHILE LOGISTIC REGRESSION CAN DETERMINE STRENGTHS OF CERTAIN PREDICTORS (beta, log odds), NEURAL NETWORKS CAN NOT
• ONCE TRAINED, ONE CAN INPUT NOVEL DATA INTO THE NETWORK WHEN THE PREDICTOR VARIABLE IS UNKNOWN AND STILL CREATE PREDICTIONS
• MULTIPLE REGRESSION TECHNIQUES ARE BASED ON STATISTICAL EFFECT SIZE ISSUES. IN ORDER TO ACHIEVE A MEANINGFUL MODEL, A CERTAIN SAMPLE SIZE IS REQUIRED
• THERE IS NO WAY TO EXAMINE OVERALL SIGNIFICANCE IN NEURAL NETWORKS
• THERE IS NO SIGNIFICANT TESTING ON PREDICTOR VARIABLES IN NEURAL NETWORKS
• THE SUCCESS OF BOTH MODELS DEPENDS ON THE AMOUNT OF DATA ALLOCATED
THE END