



AI Implications for Public Health: Equity & Inclusivity

PIPER JACKSON, PHD

THOMPSON RIVERS UNIVERSITY

Focus on Machine Learning

- Family of AI methods based on Statistics and Computing
- Derive knowledge from analysis of data
 - Rather than instructions
 - Can find patterns hard to isolate or describe by humans
- Widely applied
 - **Recommenders:** Amazon, Youtube, Facebook
 - **Prediction:** weather, stock market, wound analysis
 - **Generated content:** ChatGPT



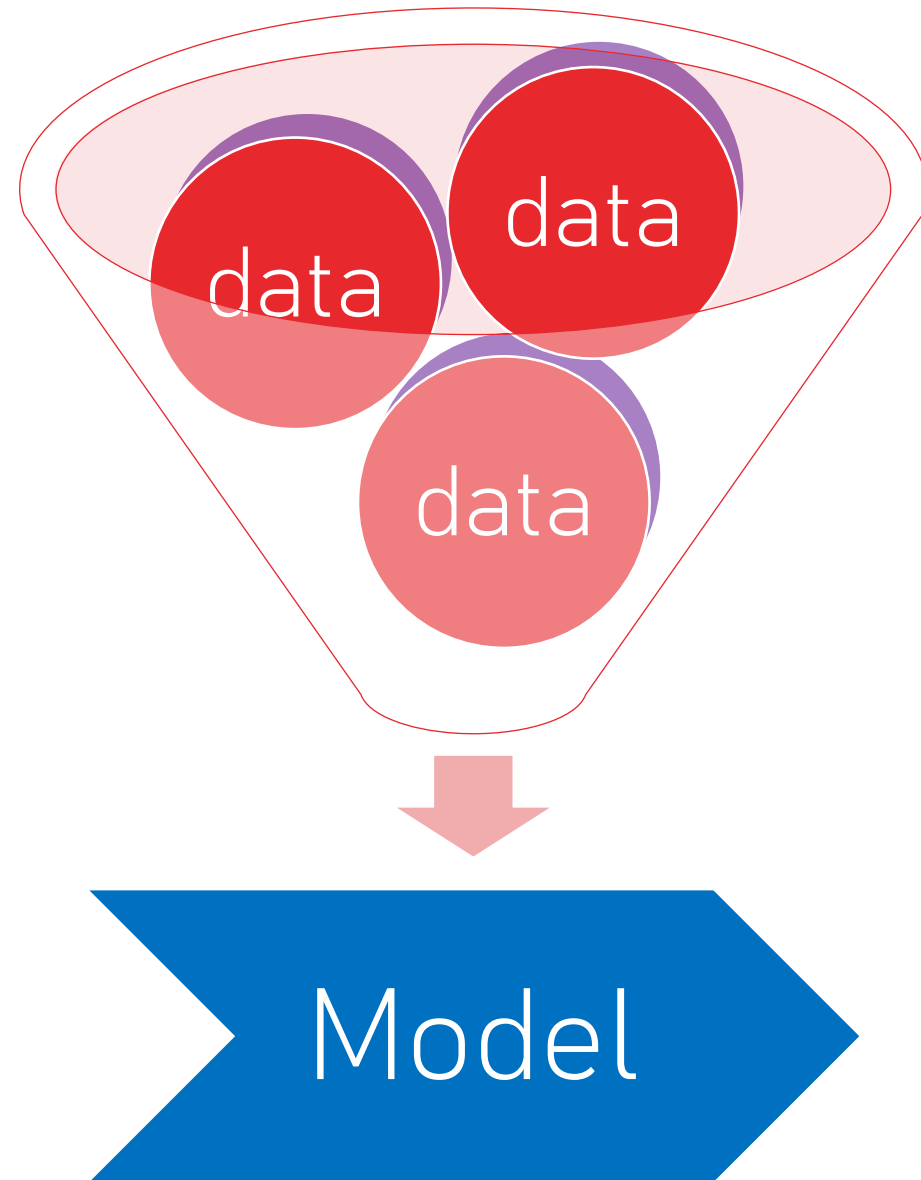
Building an ML Model

ML algorithm analyzes data and constructs a model

- Distilled representation of the data
- "Understanding" of the problem

Usually, the most computationally expensive stage

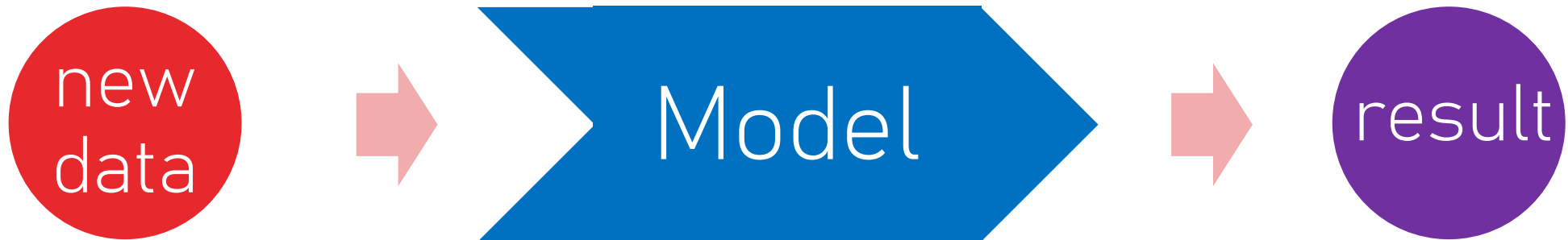
May or may not be interpretable by humans



Using an ML Model

The model can now be given new data and generate new results:

- Viewing history → recommendations
- Medical history → prediction
- Description → generated art
- Question → answer



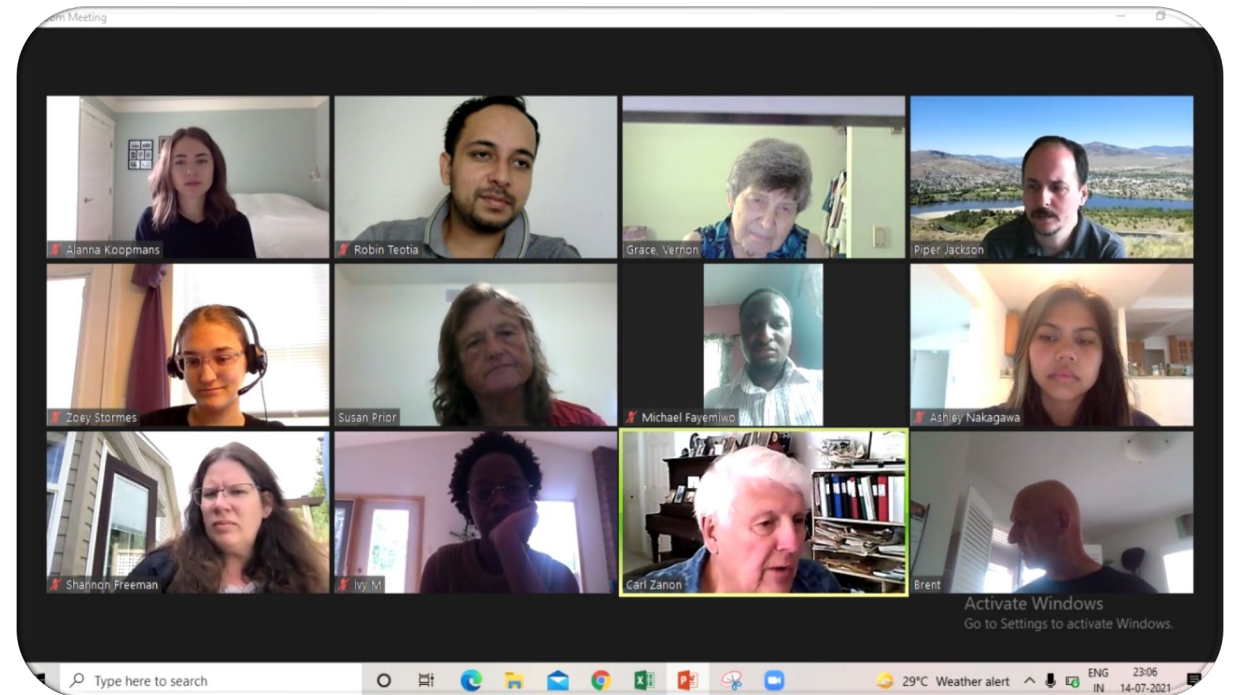
Machine Learning: Problems

- Completely dependent on data
- Missing data
 - Lack of representation
 - Missing features
- Can *bake in* inequity: *Algorithmic Bias*
 - Bias in data collection
 - Bias in data content



Machine Learning: Problems

- Completely dependent on data
- Missing data
 - Lack of representation
 - **Missing features**
- Can *bake in* inequity: *Algorithmic Bias*
 - Bias in data collection
 - Bias in data content



Patient Oriented Predictive Modelling of Healthcare Usage (POPMHU)

Machine Learning: Problems

- Completely dependent on data
- Missing data
 - Lack of representation
 - Missing features
- Can *bake in* inequity: *Algorithmic Bias*
 - Bias in data collection
 - Bias in data content



Solutions: Representation

- **Balancing data**
 - Resampling
 - Bootstrapping, oversampling
 - Generative approaches
 - SMOTE: interpolate new examples from existing records
 - GAN: deep learning generated synthetic examples to increase size of minority groups
- **Gathering the right data**
 - *Guidance on the Use of Standards for Race-Based and Indigenous Identity Data Collection and Health Reporting in Canada* (CIHI, 2022)
 - Anti-Racism Data Act (BC, 2022)





Solutions: Equity & Inclusivity

- Access to data
 - Generating synthetic data
 - Can be more freely shared and used
 - Concern: *leaking* private data
- Justice and fairness in data usage
 - *SPOR*: Strategy for Patient Oriented Research (CIHR)
 - *OCAP*[®]: Principle of *Ownership, Control, Access & Possession* of First Nations own data