Generative AI for Dairy Life Trajectory Modeling:

Transforming Milk Recording and Sensor Data into Actionable Farm Intelligence

J. Jiang, C. Maltecca, R.H. Fourdraine, and A.M. Miles*
North Carolina State University, Dairy Records Management Systems (DRMS)
*ammiles3@ncsu.edu

Background

Dairy farms collect rich, longitudinal data

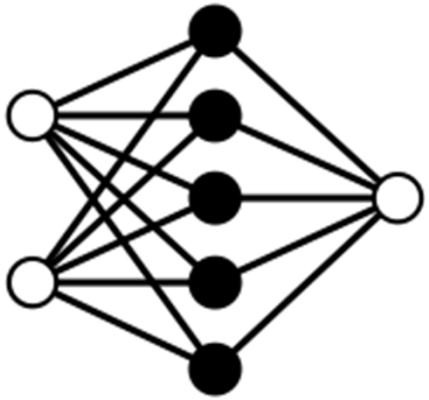
Most data is used retrospectively, after decisions have already been made

Existing models typically predict a single outcome (e.g., pregnancy or mastitis)

What if we treat a cow's life as a sequence of events and predict the next event?

(like a language model)

GPT-style Language Models



GPT = Generative Pre-Trained Transformer

Type of deep learning neural network used to generate human-like text (e.g., Chat bots)

Autoregressive prediction loops read a sequence of text and predict most probable next word

Transformer Deep Learning Architecture

Embedding

- Inputs divided into tokens
- Converted to numerical vectors

Vaswani et al 2017 “Attention is all you need”

Transformer Deep Learning Architecture

Embedding

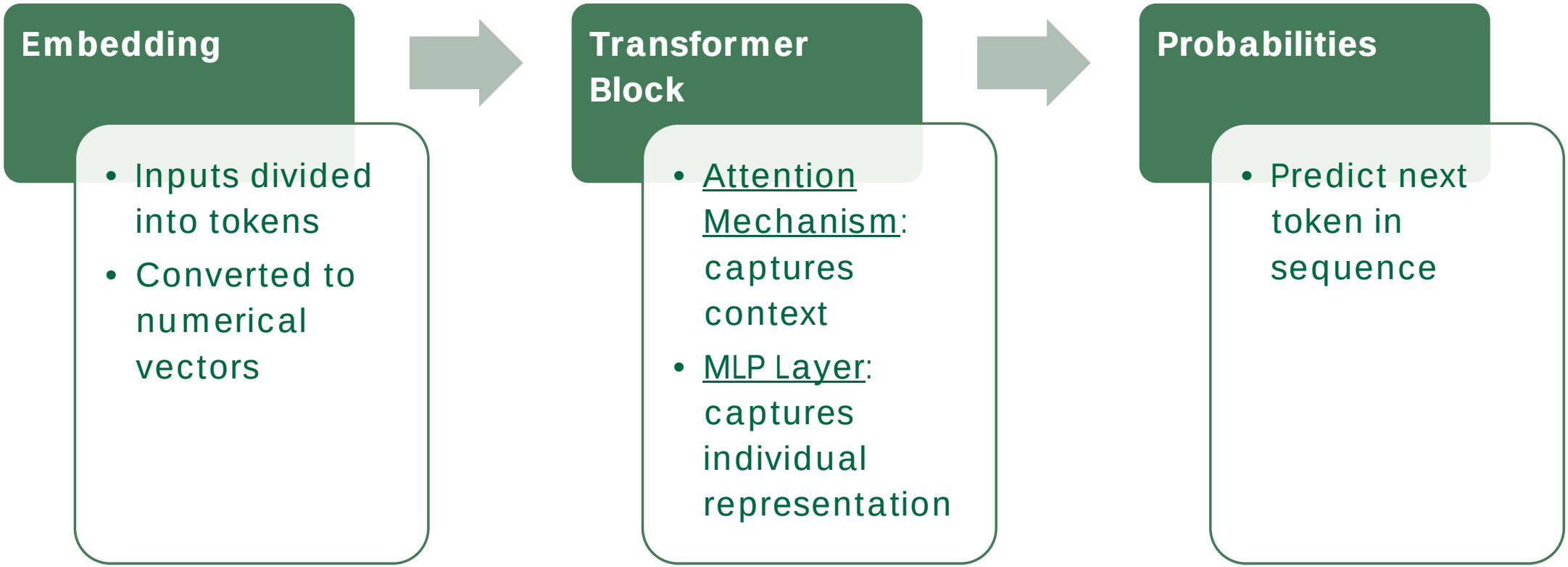
- Inputs divided into tokens
- Converted to numerical vectors

Transformer Block

- Attention Mechanism: captures context
- MLP Layer: captures individual representation

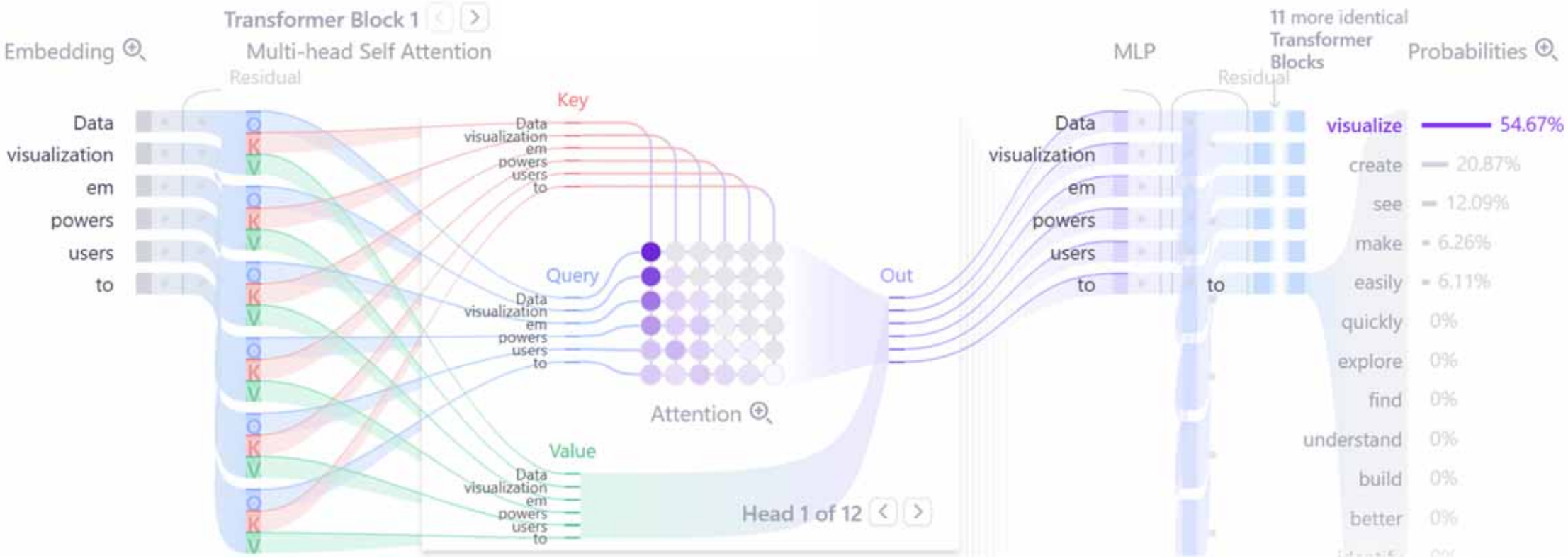
Vaswani et al 2017 "Attention is all you need"

Transformer Deep Learning Architecture



Vaswani et al 2017 "Attention is all you need"

Interactive Example (Georgia Tech)

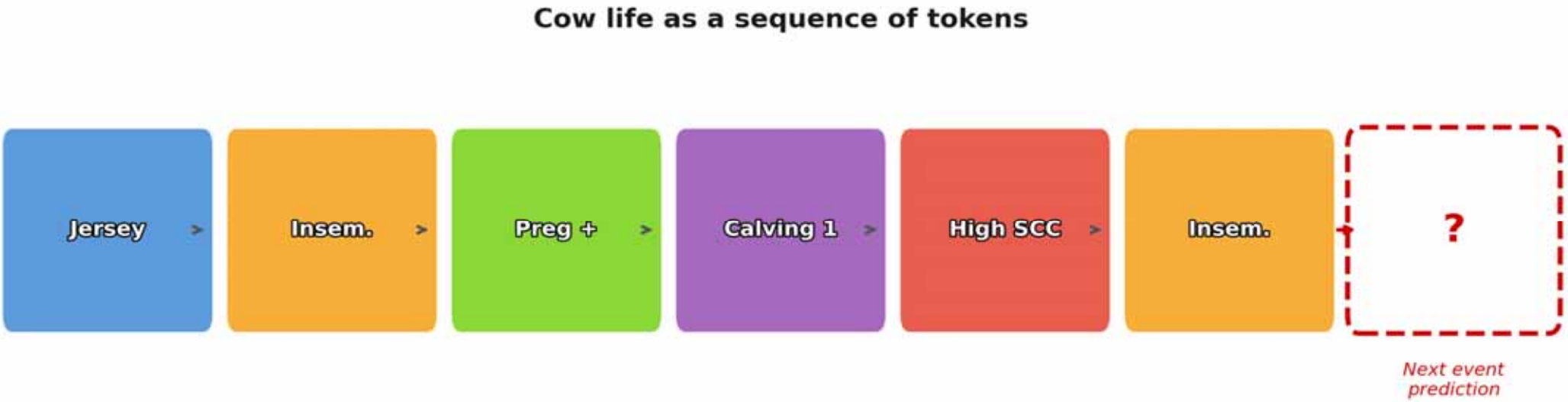


https://poloclub.github.io/transformer-explainer/

What if we treat a cow's life as a sequence of events and predict the next event?

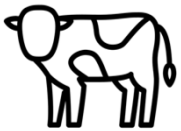
(like a language model)

Core Idea: Next Event Prediction



Token = sequence of pairs [event, age]

Proof-of-Concept Data Set

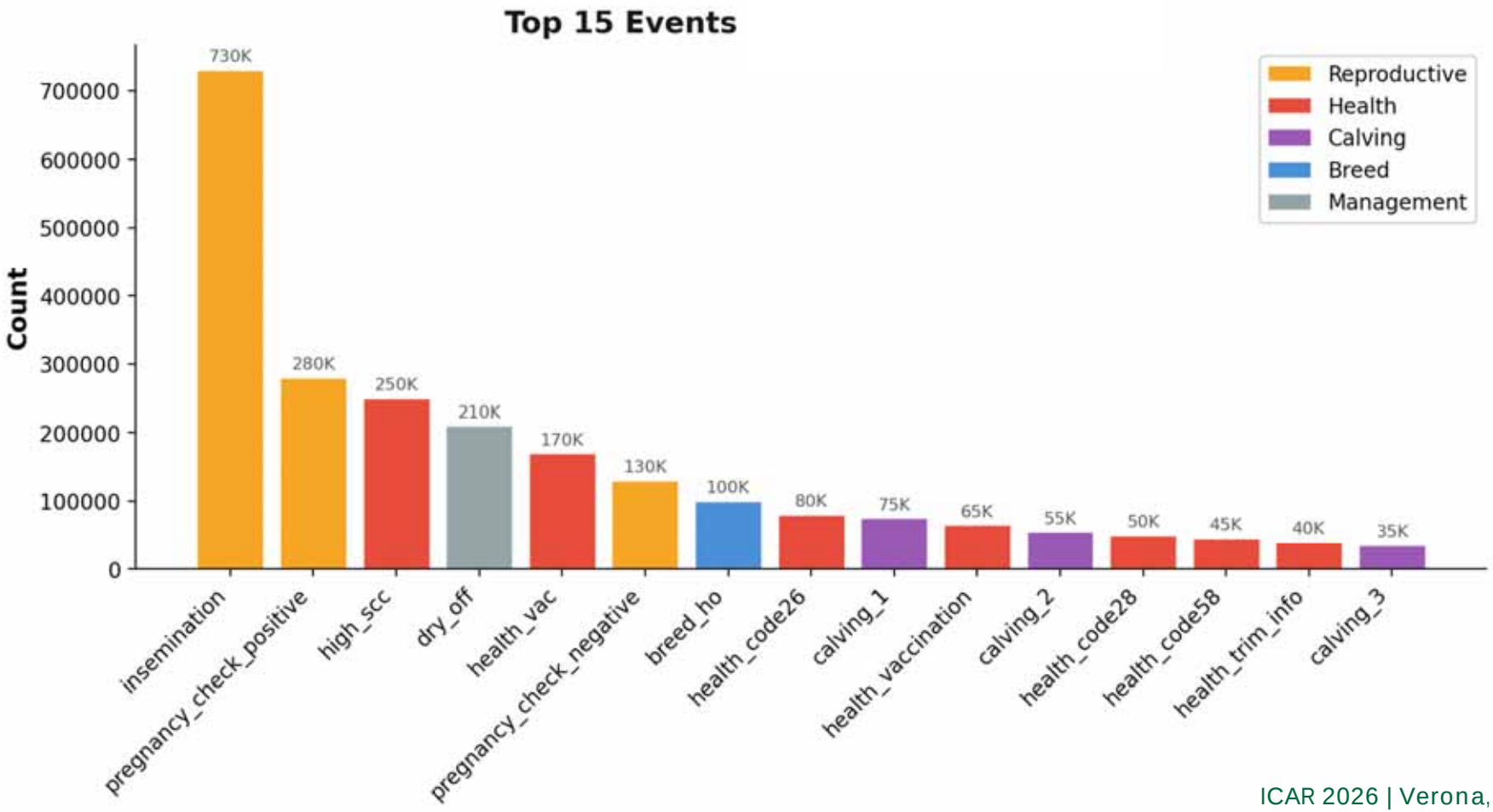


109,576 cows from 80 commercial herds

3,238,775 total events

Record types: Breed, Health, Lactation, Breeding

Proof-of-Concept Data Types



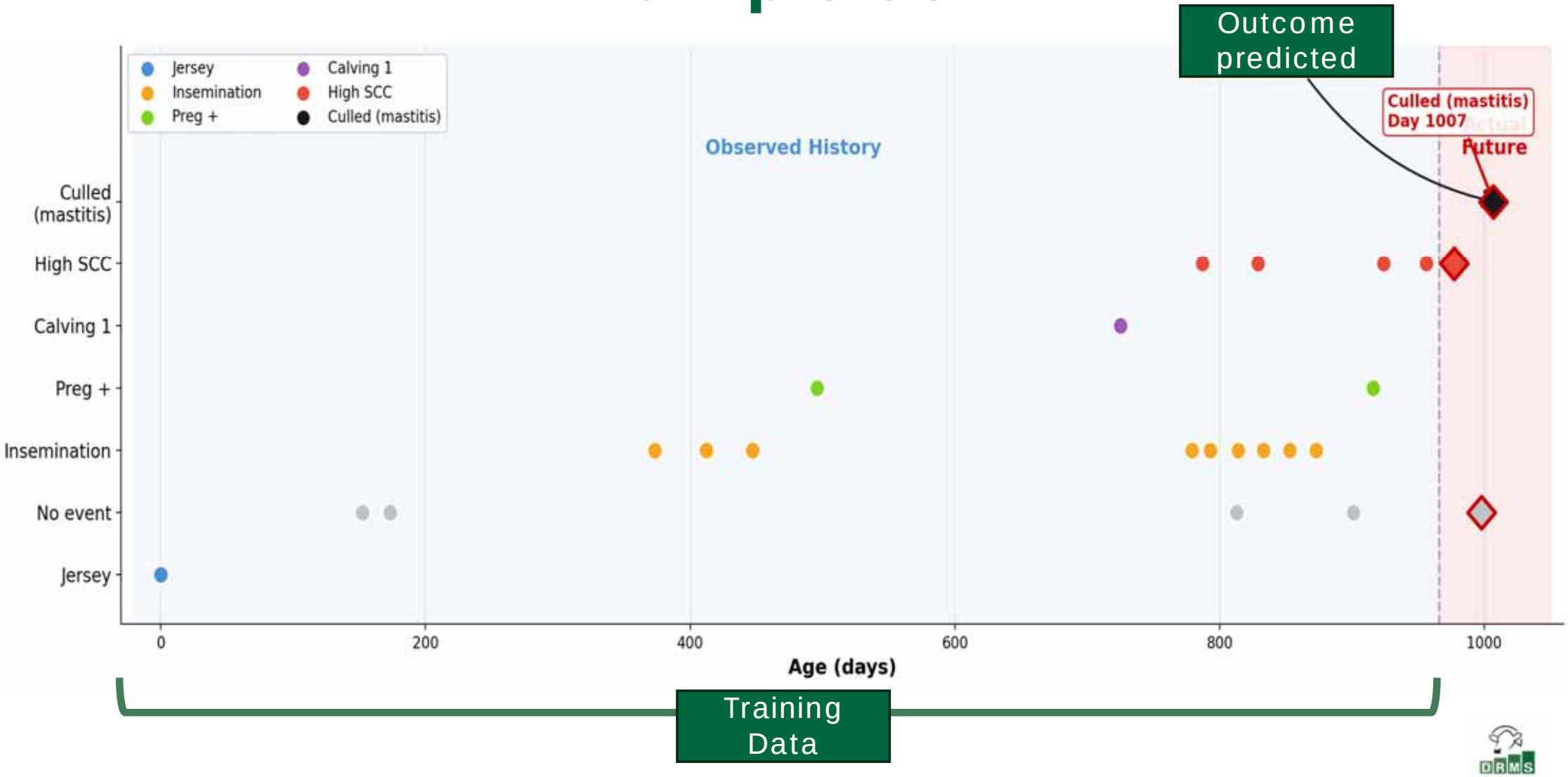
Model & Training

- GPT-style autoregressive transformer
- Dual embedding: event tokens + age
- Predicts both next event and next age

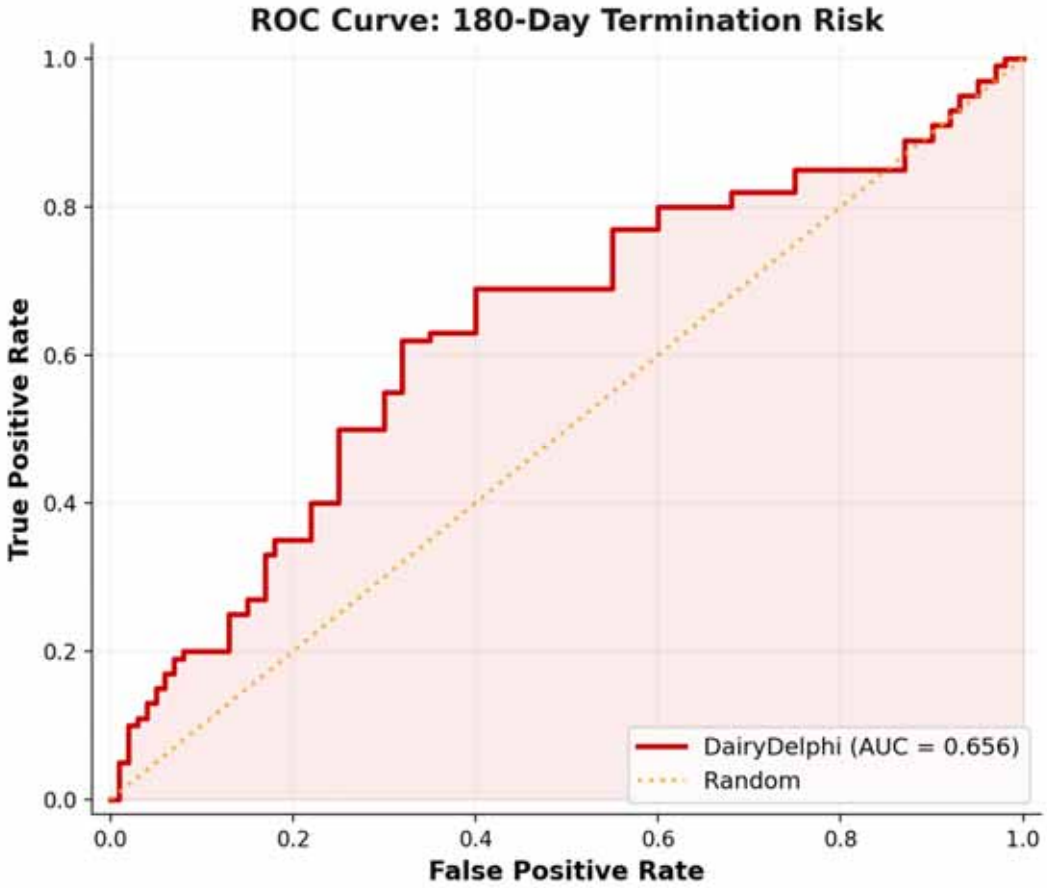
- Training time: ~5 hours (GPU)
- Hyperparameter tuning: several runs
- Inference: generate 100+ trajectories per cow

No assumptions about management rules or biological thresholds.
All relationships learned directly from observed data.

Example Cow


Outcome predicted
Culled (mastitis)
Day 1007
Actual
Future
Jersey
Insemination
Preg +
Calving 1
High SCC
Culled (mastitis)
Observed History
Culled (mastitis)
High SCC
Calving 1
Preg +
Insemination
No event
Jersey
0
200
400
600
800
1000
Age (days)
Training Data

Preliminary Evaluation



- Focus: probability of termination (culling/death) within 180 days
- Model distinguishes low-risk vs. high-risk cows
- AUC = 0.656 (initial, not yet optimized)
- Inference: ~10 hours for partial herd evaluation

Key Messages

- Transformer models need massive amounts of data for high quality predictions
- A generative model learning from 3.2M events across 109K cows can predict future survival

Next steps include:

- Improving predictive accuracy (data, architecture, tuning)
- Expand to health-specific risk scores
- Integrate genomic information and sensor data

Thank You.

Questions?

ammiles3@ncsu.edu

www.drms.org

919-661-3107