### Al in Alberta

Bruce Matichuk December 5, 2024 ESNA 2025 Economic Outlook Conference

## Overview

- 1. AI Concepts
- 2. Application
- 3. Al Basics
- 4. Agentic Systems
- 5. Trends



## About me

- MSc/PhDc in AI University of Alberta
- 30+ years in industry applying AI
- Founder
  - Celcorp Intelligent Integration
  - Poynt Local Search and Directions (Before Siri)
  - Clinitrust Secure Medical Communications Platform
  - Aidant Intelligent Technology Intelligent Recognition System
  - Health Gauge Intelligent Health Monitoring and Management Platform
  - MedWatch Non-invasive Blood Glucose Monitoring



Artificial Intelligence at the U of A

### **Ranked 2nd**

in North America for AI

### \$100 million

in AI funding since 2017

**24** Canada CIFAR AI Chairs

# WE ARE

#### the 1st Computing Science department in Canada

#### **OUR GLOBAL MISSION & PILLARS**

#### **Empower individuals to live longer, healthier lives.**

Our innovative platform will help people to proactively manage their glucose to avoid progression of diabetes and reduce the risk of other chronic conditions.



TM

11:1

### Featuring NEEDLE-FREE

CONTINUOUS GLUCOSE MONITORING

- Non-invasive ease and comfort
- Long-lasting for convenience
- Less waste, less cost

\*Not FDA Cleared. FDA submission in process

- Expandable platform for monitoring vitals
- The CGM that people not only prefer, but LOVE

#### MEDWATCH INNOVATIVE ROADMAP: AN INTELLIGENT AI-BASED HEALTH PLATFORM Empower individuals to live longer, healthier lives.





## What is Al?

- The study of systems that have human-like intelligence
- Al is mankind's most important creation
  - The power to change the world
  - The power to destroy
- Al is a mirror into ourselves
  - What is a person?
  - What does it mean to think and reason?
  - Can machines be conscious?
- Artificial Intelligence vs Machine Learning



# Demystifying AI: From Thought to Consciousness

- Turing Test
  - Can machines think?
- Strong-Al
  - An AI system can think and have mind.
- Weak-Al
  - An AI system can only act like it thinks and has a mind.
- General AI vs Narrow AI
  - AGI Artificial General Intelligence
- The AI Singularity Are we there yet?





### The Singularity Curve



### Moore's Law Curve

F



### Frontier



Currently US made Frontier runs at a peak 1.2 petaFLOPS

Note: Oracle is building a new supercomputer that will run at 2.4 zettaFLOPS

### Hardware Based Al

- NVIDIA GPU
- Google TPU
- Intel
- IBM AI chip
- <u>Neuromorphic</u> <u>Computing</u>











#### NVIDIA Grace Hopper Superchip





### Cerebras Wafer Scale Engine (WSE-3)

- The largest chip ever built
- 4 Trillion transistors
- 900,000 AI optimized compute cores.
- 44 GB on-chip memory
- 125 Petaflops
- External memory: 1.5TB, 12TB, or 1.2PB
- Trains AI models up to 24 trillion parameters

### The AI Singularity Curve

Floating-point operations per second (FLOPS) 1,000,000,000,000,000,000,000,000,000 100,000,000,000,000,000,000,000,000 10,000,000,000,000,000,000,000,000 1,000,000,000,000,000,000,000,000 100,000,000,000,000,000,000,000,000 10,000,000,000,000,000,000,000,000 1,000,000,000,000,000,000,000,000 100,000,000,000,000,000,000,000 10,000,000,000,000,000,000,000 1,000,000,000,000,000,000,000 100,000,000,000,000,000,000 10,000,000,000,000,000,000 1,000,000,000,000,000,000 100,000,000,000,000,000 10,000,000,000,000,000 1,000,000,000,000,000 100,000,000,000,000 10,000,000,000,000 1,000,000,000,000 100,000,000,000 10,000,000,000 1,000,000,000 100,000,000 10,000,000 1,000,000 100,000 10,000 1,000 100 10 0



- Exponential growth of supercomputing power, 1995-2060 (logarithmic scale)
- 2025 Human brain
  - 100 billion neurons
- 2036 1000 Brains
- 2047 1 million brains

### **Artificial Intelligence Market Size**

#### **Global Artificial Intelligence Market** On-Premise Cloud Size, By Deployment Mode, 2024-2033 (USD Billion) 4000 3,527.8 3500 3000 2,707.5 2500 2,077.9 2000 1,594.7 1500 1,223.8 939.3 720.8 1000 553.2 424.6 325.8 250.1 500 2023 2024 2025 2027 2028 2029 2031 2032 2033 2026 2030

The Market will Grow 30.3% The Forecasted Market \$3,527.8B I market.us Size for 2033 in USD:

## AI Topics

- Symbolic Processing
- Machine Learning
  - Neural Nets
  - Deep Learning
  - Reinforcement Learning
- Search/Games
- Generative Al
- Attention/Transformer Architecture
- Large Language Models ChatGPT40
- Diffusion DALLE 2
- Intelligent System Orchestration



#### 

## Symbolic Reasoning

Brothers are siblings  $\forall x, y \; Brother(x, y) \Rightarrow Sibling(x, y).$ "Sibling" is symmetric  $\forall x, y \; Sibling(x, y) \Leftrightarrow Sibling(y, x).$ One's mother is one's female parent  $\forall x, y \; Mother(x, y) \Leftrightarrow (Female(x) \land Parent(x, y)).$ A first cousin is a child of a parent's sibling

 $\begin{array}{lll} \forall x,y \ \ FirstCousin(x,y) \ \Leftrightarrow \ \exists \, p,ps \ \ Parent(p,x) \land Sibling(ps,p) \land \\ Parent(ps,y) \end{array}$ 

### **Supervised Learning**



### **Classification vs Regression**



## Clustering



### Examples of symbolic reasoning

- Database queries
- Constraint Systems
- Planners and Schedulers
- Formal Verification Tools
- Temporal Reasoning
- Description Logics

## Machine Learning

- Supervised Learning
  - Classification
  - Regression

Ę

- Unsupervised Learning
  - Clustering
  - Dimensionality Reduction
  - Embeddings
  - Feature learning
- Reinforcement Learning (RL)

### Reinforcement Learning



## Q-Learning



## **Reinforcement Learning**

- Value-Based Methods
  - Q-learning
- Policy-Based Methods
  - Policy gradients
- Model-Based Methods
  - Example: model of the environment
- Applications
  - Robotics
  - Game playing (e.g., AlphaGo)
  - Autonomous driving.



## **Artificial Neural Nets**





## **Deep Learning**

- Size matters
- 100s of layers



### Backpropagation



### Backpropagation



### Backpropagation



### **Deep Learning**

- Regular Neural Networks have 1 or 2 hidden layers
- Each layer allows more complexity in the decision making
- Deep Neural Networks use many layers
  - AlphaGo uses 13 layers for each decision (value and policy)
  - Combined with tree search
    - Effective layers = thousands
  - Note: AlphaGo in 1 year advanced automated Go by 20 years
- Microsoft Deep Learning net for images
  - 152 layers
- ChatGPT 500 + layers

### **Hierarchical Processing**



## **Deep Learning**

- Size matters
- 100s of layers



## **Deep Learning**

- Regular Neural Networks have 1 or 2 hidden layers
- Each layer allows more complexity in the decision making
- Deep Neural Networks use many layers
  - AlphaGo uses 13 layers for each decision (value and policy)
  - Combined with tree search
    - Effective layers = thousands
  - Note: AlphaGo in 1 year advanced automated Go by 20 years
- Microsoft Deep Learning net for images
  - 152 layers
- ChatGPT 500 + layers

### **Hierarchical Processing**



## Transformers and ChatGPT

- Based on Attention
- Large Language Model
- Super Intelligent
- 1.2 Billion users
- Fastest growing application
- More progress than predicted
- Emergent Behavior
- Feared
- Surprisingly useful for many tasks
- Rapidly changing field
- API Available from various vendors
- Plugins
- Fine-tuning



# The Transformer System

- The DNA of Super-Intelligence
  - Small code footprint
  - Massive Parameter Size
  - Trained on all of mankind's knowledge
  - Black Box
- Generative Al



### Generative Text – Anything you can do...

- GPT (Generative Pre-trained Transformer)
  - Models like GPT-3 and GPT-4, developed by OpenAI, can generate human-like text based on prompts, create articles, answer questions, and even write code.
- Chatbots
  - Advanced conversational agents like ChatGPT use generative AI to engage in natural, human-like conversations.
- Text Summarization
  - Tools that condense long pieces of text into shorter summaries while preserving key information.
- Creative Writing
  - Al systems that can write stories, poems, and other forms of creative literature.

### Generative AI - Images

- DALLE
- MidJourney
- Stable Diffusion

### **Generative AI Images**

#### **Diffusion Networks**

- Al models that can generate arbitrary images from textual descriptions.
- DALL-E, Midjourney, Stable Diffusion

GANs (Generative Adversarial Networks)

• Used to create realistic images, art, and even deepfakes by pitting two neural networks against each other.

#### Style Transfer

• Al can transform the style of an image, such as making a photo look like a painting by a famous artist.



## DALLE

an armchair in the shape of an avocado"













### **Generative AI Music and Sound**



## **Generative Al Video**

### Conversational System

- Based on LLMs
- Require an Orchestration System
  - Determine topic
  - Map conversation to workflow
  - Specialized pre-defined connections to many systems
- Managed Prompting
- Can manage context for long term memory



### How does the solution work?



#### **Conceptual Diagram**



### **Example Execution of a Workflow**

Query: How did my portfolio compare against S&P and Russell 2000 in the last 3 years?



### Intelligent Process with Orchestration



### Modern AI Miracles

- Natural Language Understanding & Generation
- Image Generation & Style Transfer
- Autonomous Vehicles
- Protein Structure Prediction
- Personalized Healthcare
- Real-Time Language Translation
- Deep Reinforcement Learning
  - AlphaGo, AlphaFold
- Hyper-Realistic Digital Assistants
  - Alexa, Siri
- Content Creation
- Video Creation
- Game Creation

- Neural Artistry
  - Dalle, MidJourney
- Emotion Recognition
- Voice Cloning & Speech Synthesis
  - Deep Fake
- Al Companions
- Al-Powered Drug Discovery
- Autonomous Industrial Robotics
- Advanced Human-Machine Interfaces
- Personalized Education
- Synthetic Biology and DNA Design
- Automatic Programming

### Modern AI Applications: Financial Industry

- Algorithmic Trading
- Fraud Detection and Prevention
- Credit Risk Assessment
- Personalized Financial Services
- Chatbots and Virtual Assistants
- Sentiment Analysis for Market Prediction
- AI-Powered Investment Advisors (Robo-Advisors)
- Customer Behavior Analytics
- Regulatory Technology (RegTech)
- Predictive Analytics for Loan Default

- Automated Underwriting
- Natural Language Processing for News Impact Analysis
- Portfolio Optimization
- Credit Scoring for Underbanked Populations
- Market Anomaly Detection
- Know Your Customer (KYC) Automation
- Risk Management through Deep Learning
- Fraud Detection in Real-Time Payment Systems
- Anti-Money Laundering (AML) Compliance Monitoring
- Insurance Claims Processing Automation

### Trends

- New models with superhuman abilities
  - OpenAI, Google, AWS
- Multimodal Models
- Real Time Simulated Worlds
- Agentic Systems
  - Mixture of Agents
- Goal Directed Behavior

