Artificial Intelligence: From Mathematical Fitting to Human-like Intelligence

AI Seminar Series

January 2026

Contents

Part 1 Classic Machine Learning - Manual Feature Extraction (15 mins)

Regression Models: The Mathematical Foundation

- **Simple Linear Regression**: 建立 $y = wx + b$ 的初步模型。探讨最小二乘法（OLS）如何最小化残差平方和。
- **Polynomial Regression**: 处理非线性特征的曲线拟合。引入基函数扩展，讨论阶数过高导致的过拟合风险。
- **Logistic Regression**: 分类任务的基石。通过 Sigmoid 函数将线性输出映射至 $[0, 1]$ 概率空间，引入交叉熵损失函数。

Statistics-based Classification

- **K-Nearest Neighbors (KNN)**: 基于欧氏距离或曼哈顿距离的朴素投票法。讨论 K 值选取对决策边界平滑度的影响。
- **Support Vector Machines (SVM)**: 寻找最大化分类间隔（Margin）的超平面。探讨核技巧（Kernel Trick）如何将数据映射至高维空间以解决线性不可分问题。
- **Decision Trees**: 基于信息增益（ID3）或基尼系数（CART）的递归分裂。模拟人类决策逻辑，是随机森林与 GBDT 的基础单元。
- **Naive Bayes**: 基于条件概率独立性假设。在文本分类和垃圾邮件过滤中展现出极高的计算效率。

Part 2 Neural Architecture - The Structural Build (20 mins)

The Artificial Neuron

- **Biological Inspiration**: 模拟树突（输入）、胞体（加权求和）与轴突（输出）的信号传递过程。
- **Weights and Biases**: 参数化的核心。权重代表特征重要性，偏置提供模型激活的灵活性（位移系数）。
- **Activation Functions**: 引入非线性。对比 ReLU（解决梯度弥散）、Sigmoid（概率映射）与 Tanh（零中心化）的应用场景。

Multi-Layer Perceptron (MLP)

- **Input, Hidden & Output Layers**: 数据的层级表征。隐藏层作为“黑盒”负责自动提取特征，消除对人工特征工程的依赖。
- **Full Connectivity**: 全连接层。神经元之间的稠密连接确保了信息的最大化交换，但也带来了参数量过大的问题。
- **Universal Approximation Theorem**: 证明了单隐藏层非线性 MLP 只要宽度足够，即可拟合任意连续函数。

Part 3 The Engine of Learning - Optimization & Backpropagation (20 mins)

The Concept of Error

- **Loss Functions**: 定义 MSE（回归）与 Cross-Entropy（分类）。讨论损失函数作为模型优化“导航标”的数学本质。
- **Cost Function Visualization**: 理解高维损失曲面中的局部最小值、全局最小值与鞍点。

Gradient Descent: The Navigator

- **Gradient Concept**: 损失函数对参数的偏导数向量，指向下降最快的方向。
- **GD, SGD & Mini-batch GD**: 对比全量梯度下降、随机梯度下降与小批量梯度下降在收敛速度与内存占用上的平衡。
- **Learning Rate**: 探讨超参数调整。过大导致震荡，过小导致停滞。引入学习率衰减（Decay）策略。

Backpropagation: The Knowledge Correction

- **The Chain Rule**: 梯度传导的数学基石。通过复合函数求导将误差逐层反向传播。
- **Backward Pass**: 计算每层权重对总误差的贡献。
- **The Vanishing Gradient Problem**: 探讨随着深度增加，梯度在乘法链中趋近于零的现象，以及它如何制约了早期深层网络的发展。

Part 4 Computer Vision - Capturing Spatial Patterns (20 mins)

Convolutional Foundations

- **Digital Pixels & Kernels**: 机器通过滑动窗口（卷积核）进行局部特征感知。
- **Convolutional Layers**: 权重共享与局部连接。自动提取边缘、纹理、形状等层级化视觉特征。
- **Pooling**: 空间降采样。增大感受野，提供平移不变性并减少计算负载。

The Evolution of Vision Models

- **LeNet & AlexNet**: 从邮政识别到 2012 ImageNet 夺冠。探讨 GPU 算力如何激活了沉睡的神经网络理论。
- **VGGNet**: 强调 3x3 小卷积核的反复堆叠。模块化结构使得网络深度首次突破 10 层。
- **ResNet**: 引入 Shortcut Connection（残差连接）。通过恒等映射解决了深层网络的退化问题，将深度推向千层。

Part 5 Sequence and Language - Modeling the Arrow of Time (20 mins)

Recurrent Architectures

- **RNN**: 时间步之间的权值共享。通过隐藏状态记录序列历史。
- **LSTM & GRU**: 门控机制（输入门、遗忘门、输出门）。解决了长序列训练中的长期依赖遗忘问题。

NLP Basics

- **Tokenization**: 从 Character-level 到 Byte-Pair Encoding (BPE)。探讨 AI 预处理语言的逻辑。
- **Word2Vec**: 分布式表征。通过上下文预测（CBOW/Skip-gram）将语义映射为连续向量。
- **Seq2Seq**: Encoder-Decoder 范式。引入上下文向量（Context Vector）实现变长序列转换。

Part 6 The Transformer Revolution - The Great Unification (20 mins)

Attention is All You Need

- **Self-Attention**: 抛弃顺序依赖。通过 Q, K, V 矩阵计算全局关联性，支持超大规模并行计算。
- **Multi-Head Attention**: 在多个子空间内学习特征，同时关注语法、语义与逻辑关系。
- **Positional Encoding**: 弥补 Transformer 缺乏位置顺序感的缺陷。

Pre-training Paradigm

- **BERT**: 双向理解。通过 Masked LM 和下句预测学习通用语义特征。
- **GPT Series**: 预训练生成的暴力美学。坚持 Decoder-only 架构，揭示了预测下一个 Token 即可产生智能。

Cross-modal: ViT

- **Vision Transformer**: 将图像切片视为“词块”。证明了视觉任务不再需要专门的卷积结构，统一架构时代开启。

Part 7 Generative AI and Large Models - The New Frontier (20 mins)

Generative Paradigms

- **GAN**: 博弈论的应用。生成器与判别器在对抗中进化，产出高保真伪造图像。
- **Diffusion Models**: 基于非平衡热力学的逆过程。从噪声中“洗”出图像，Midjourney 与 Stable Diffusion 的底层逻辑。

LLM Anatomy

- **Scaling Laws**: 讨论参数量、算力与数据规模之间的幂律关系，以及“涌现”能力的起源。
- **MoE (DeepSeek V3)**: 探讨稀疏激活。通过分发器（Router）仅激活部分参数，极大提升推理能效比。

Alignment, Reasoning & Agents

- **RLHF & RAG**: 如何通过反馈对齐价值观，以及利用外部检索解决“幻觉”问题。
- **Chain of Thought (CoT)**: 逐步推理。讨论慢思考（System 2）如何提升大模型的数学与逻辑能力。
- **Agents**: AI 智能体。从“对话”转向“行动”，具备任务规划、工具调用与自主决策能力。

Local Intelligence

- **Quantization**: 讨论 4-bit 量化。降低 VRAM 占用，使个人显卡跑起千亿参数模型。
- **Path to AGI**: 展望通用人工智能。讨论具身智能（Embodied AI）与世界模型的结合。