**Discorso per presentazione (10 minuti)**

**Slide 1: Titolo (15 secondi)**

**Slide 2: Obiettivi della Tesi (45 secondi)**

• Studio teorico: Analisi approfondita dei fondamenti matematici e delle architetture dei quattro modelli

• Analisi metodologica: Valutazione dell'efficacia in scenari reali attraverso tre casi di studio diversificati

• Addestramento e valutazione: Tre casi studio per tipologia di problema - regressione e classificazione - e natura dei dati: tabellari, serie storiche e immagini

• Studio di ablazione post-training: Analisi dell'impatto del pruning per misurare il compromesso tra complessità e prestazioni

**Slide 3-4: MLP e Backpropagation (30 secondi)**

• MLP: rete neurale con strati input, nascosti e output che elabora tramite somma ponderata e funzioni di attivazione

• Backpropagation: algoritmo di apprendimento che propaga l'errore all'indietro, utilizzando chain rule e gradient descent per aggiornare pesi e bias

**Slide 5: Curve di Bézier - FOCUS TEORICO (45 secondi)**

• Fondamento matematico chiave: Le curve di Bézier sono curve parametriche dove tutte le coordinate dipendono dalla variabile t

• Formula lineare e quadratica

**Slide 6-7: Dalle Bézier alle B-Spline - FOCUS TEORICO (1.5 minuti)**

• B-Spline: Funzioni polinomiali a tratti che definiscono le attivazioni KAN in modo flessibile

• Proprietà fondamentali:

o Continue e differenziabili

o Controllo locale: modificare un punto influisce solo localmente

o Parametri addestrabili tramite backpropagation

• Algoritmo de Boor-Cox: Metodo ricorsivo per costruzione B-spline

• Griglia di controllo: Insieme dei punti che definiscono la forma della B-spline

**Slide 8-9: Teoremi Universali - FOCUS TEORICO (1 minuto)**

• Universal Approximation Theorem (UAT): Base teorica delle MLP

• Kolmogorov-Arnold Representation Theorem (KART): Base teorica delle KAN

**Slide 10-11-12: Kolmogorov-Arnold Network - FOCUS PRINCIPALE (1.5 minuti)**

• Innovazione chiave: Invece di pesi scalari apprendibili, le KAN hanno funzioni univariate apprendibili

• Architettura: Esempio

• Differenza fondamentale: MLP alternano operazioni lineari e non-lineari, KAN sostituiscono ogni prodotto peso-input con funzioni parametriche

• Vantaggio: Superano la curse of dimensionality - l'errore dipende dalla risoluzione spline, non dalla dimensionalità input

• Dalle Spline: Controllo locale, utili per continual learning, possono essere rese sparse

• Dalle MLP: Backpropagation, composizionalità, scalabilità

• Risultato: Scaling laws più favorevoli rispetto alle MLP tradizionali

**Slide 13-14: Altri Modelli (30 secondi)**

• XGBoost: Ensemble sequenziale

• Random Forest: Ensemble parallelo

• CNN: Architettura modificata sostituendo classificatori fully-connected con KAN

**Slide 15-16: Metodologia HPO (30 secondi)**

• Random Search: Scelto per efficienza, scalabilità e parallelizzazione

• Tecnologie: PyTorch, pykan, scikit-learn, XGBoost su cluster HPC Bologna

**Slide 17: Metriche (15 secondi)**

• Metriche standardizzate per regressione e classificazione con intervalli di confidenza bootstrap

**Slide 18-22: Caso 2 PM2.5 - FOCUS CASO STUDIO (1 minuto)**

• Problema: Classificazione inquinamento PM2.5 in 6 classi AQI da 453 città indiane (2010-2023)

• Features: 24 variabili con lag features temporali

• Risultati

• Studio di ablazione: KAN mantiene 0.7710 con 70% pruning e 3.0× compressione

**Slide 23-25: Caso 3 Immagini - FOCUS CASO STUDIO (1 minuto)**

• Problema: Classificazione età da volti UTKFace in 4 fasce

• Approccio: CNN + classificatore finale (MLP vs KAN)

• Risultati

**Slide 26: Conclusioni (30 secondi)**

• Messaggio principale: Non esiste soluzione universale - dipende da problema e vincoli

• KAN: Non sostituto universale MLP, ma alternativa valida e promettente dove leggerezza modello è cruciale