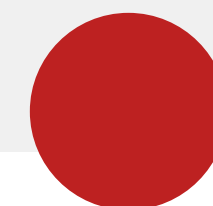


Davide DONADIO

LLM a progettazione di *LUNAR:*

un linguaggio a dominio specifico per giochi di carte collezionabili



「AGENDA」

01

MAGIC: THE GATHERING

02

LLM

03

LUNAR VS FORGESCRIPT

04

ANALISI ESPLORATIVA

05

ESPERIMENTO

06

RISULTATI

07

CONCLUSIONI

01. Una carta di Magic

- Nome della carta
- Layout
- Costo di mana
- Tipo, supertipo e sottotipo
- **Effetto**
- Note legali e artista
- forza / costituzione - Lealtà



● 02. LLM

I **Large Language Models** (LLM) sono potenti modelli di apprendimento profondo addestrati su **enormi set di dati** testuali. Il loro nucleo, è composto da encoder e decoder, ognuno con capacità di **auto-attenzione**, che permettono di estrarre significati e relazioni da sequenze di testo.

Sono flessibili e possono eseguire diversi compiti:

- rispondere a domande
- riassumere documenti
- tradurre lingue
- completare frasi
- trasformare e creare contenuti



03. Lunar v ForgeScript



ForgeScript

Name:Lightning Helix
ManaCost:R W
Types:Instant
A:SP\$ DealDamage | Cost\$ R W |
ValidTgts\$ Any | NumDmg\$ 3 |
SubAbility\$ DBGainLife |
SpellDescription\$ CARDNAME
deals 3 damage to any target and
you gain 3 life.
SVar:DBGainLife:DB\$ GainLife |
LifeAmount\$ 3
Oracle:Lightning Helix deals 3
damage to any target and you
gain 3 life.

Lunar

name: Lightning Helix
mana_cost: R W
card_type: instant
effects:
 effect:
 type: base
 mode: damage
 target: any_target
 amount: 3
 effect:
 type: base
 mode: life_gain
 target: card_owner
 amount: 3
oracle_text: <Lightning Helix deals 3
damage to any target and you gain 3
life.>



35K

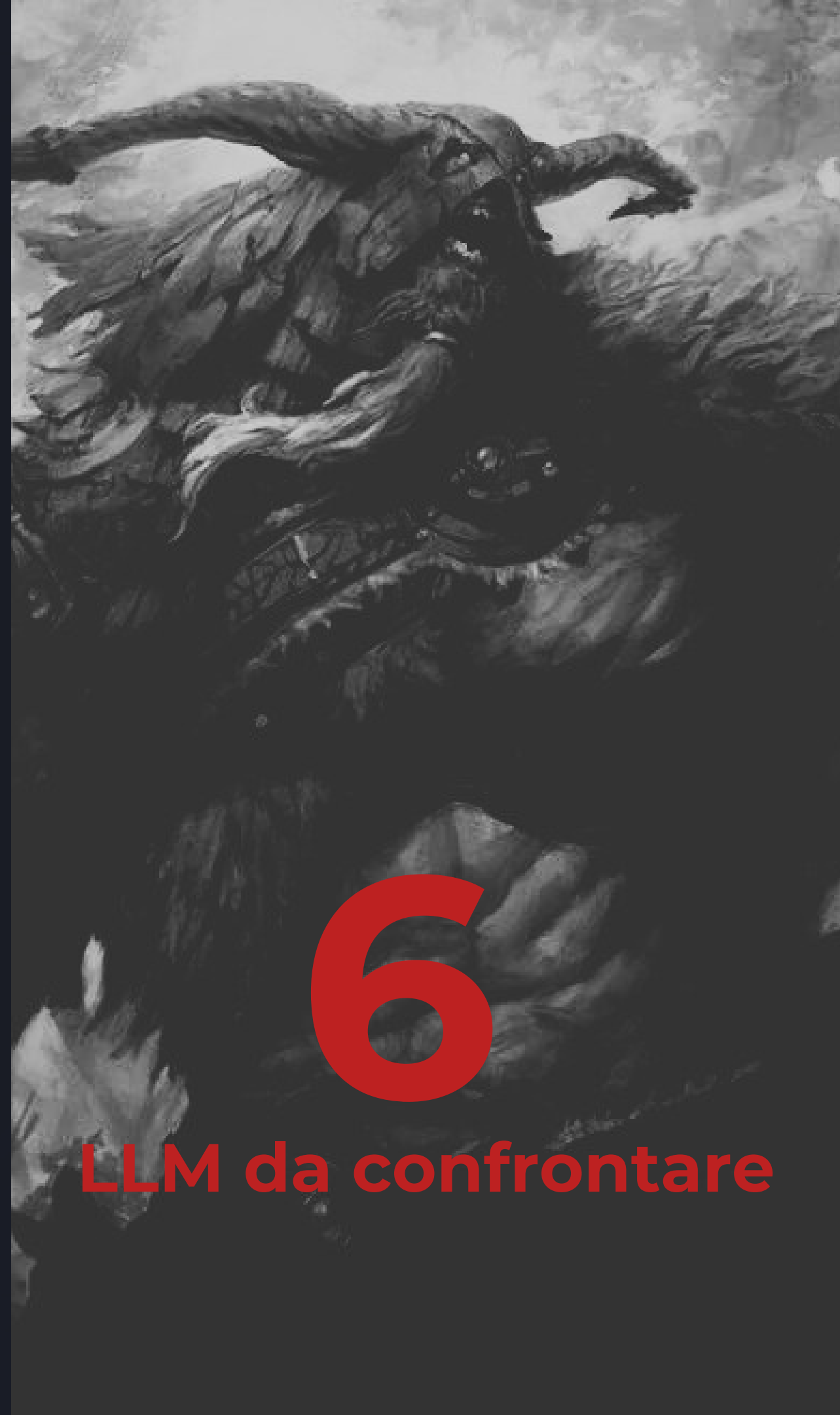
Carte analizzate



20K

Effetti Unici

CYNTHIA SHEPPARD



6

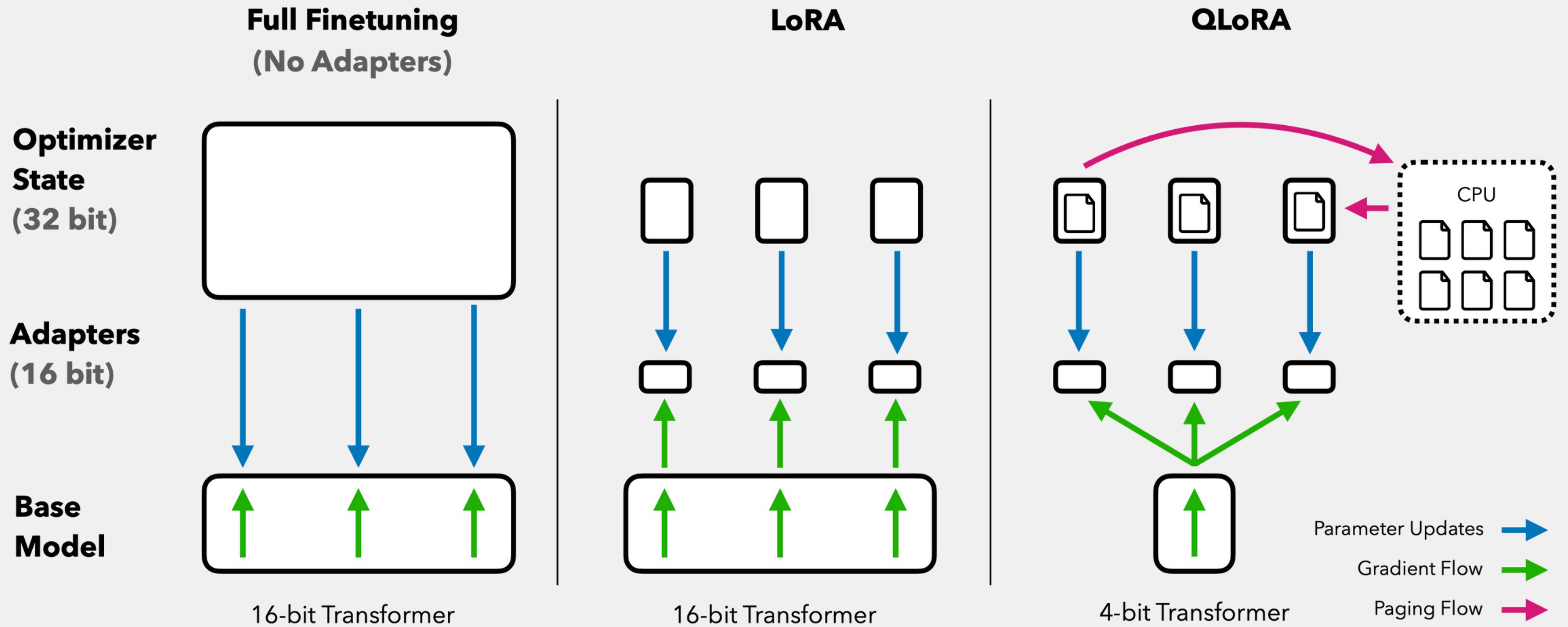
LLM da confrontare

● 05. Autotrain Advanced

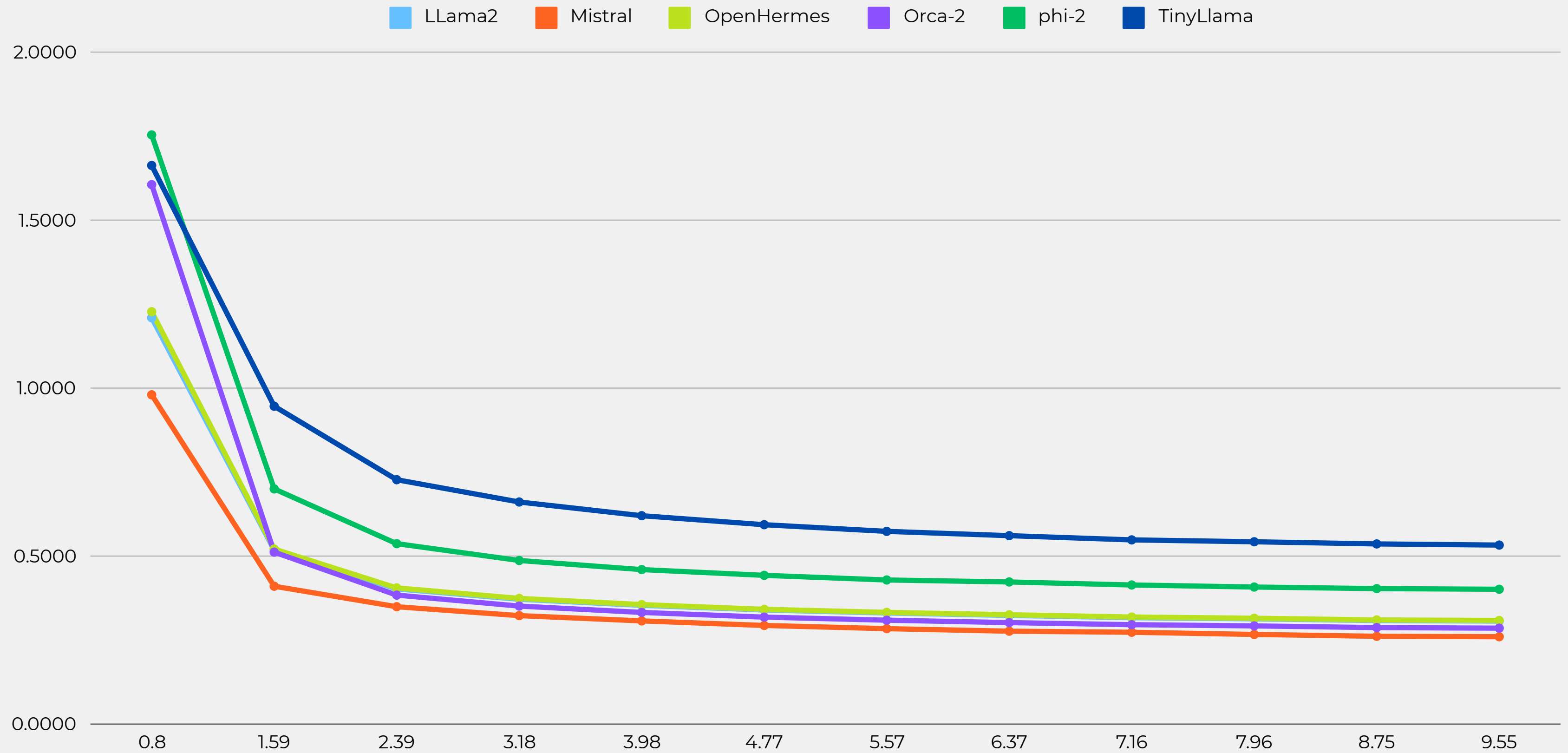


- Strumento per addestrare svariati tipi di modello
- Usa metodi di addestramento allo stato dell'arte
- PEFT — LoRA — **QLoRA**

05. QLoRA

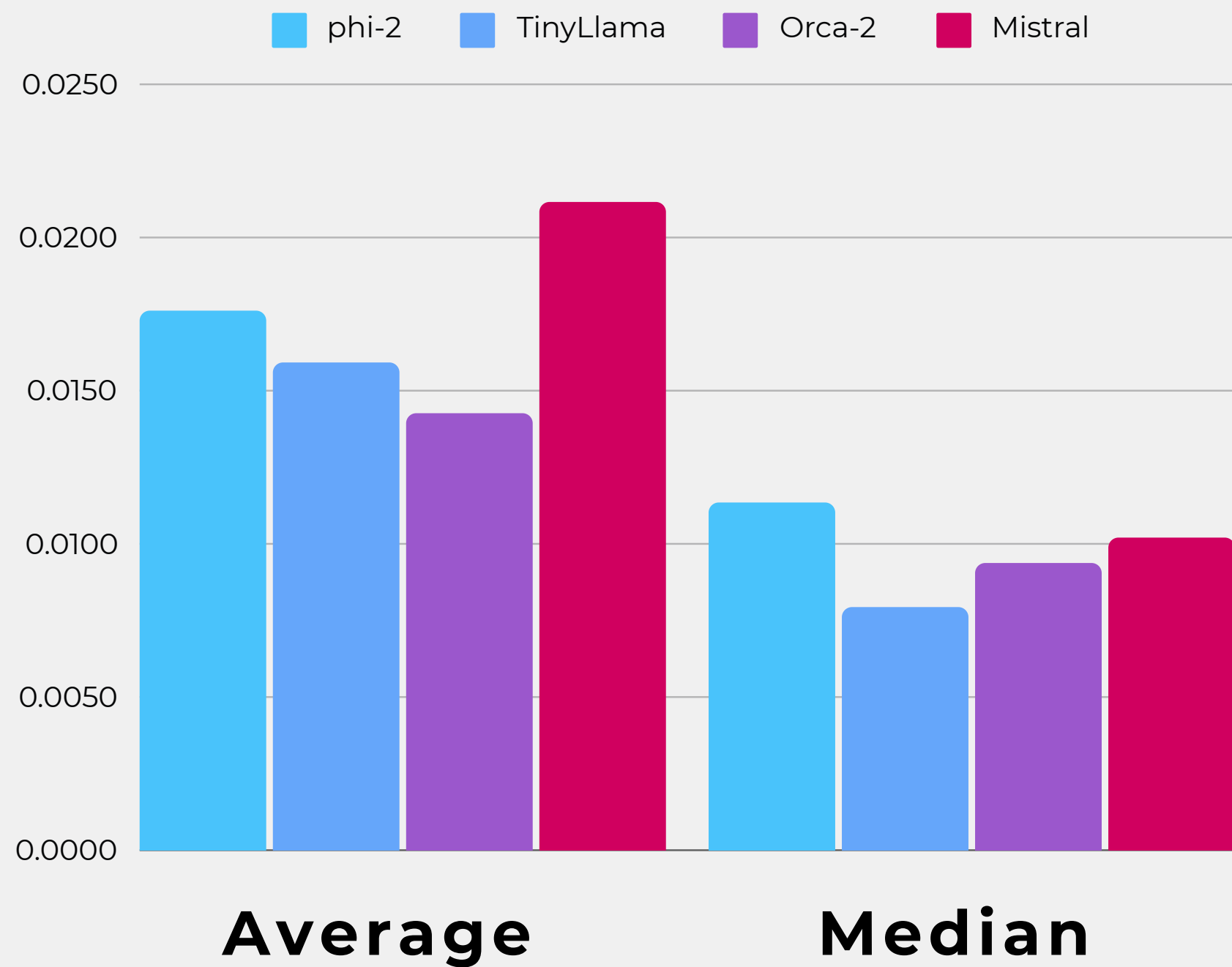


05. Train loss trend

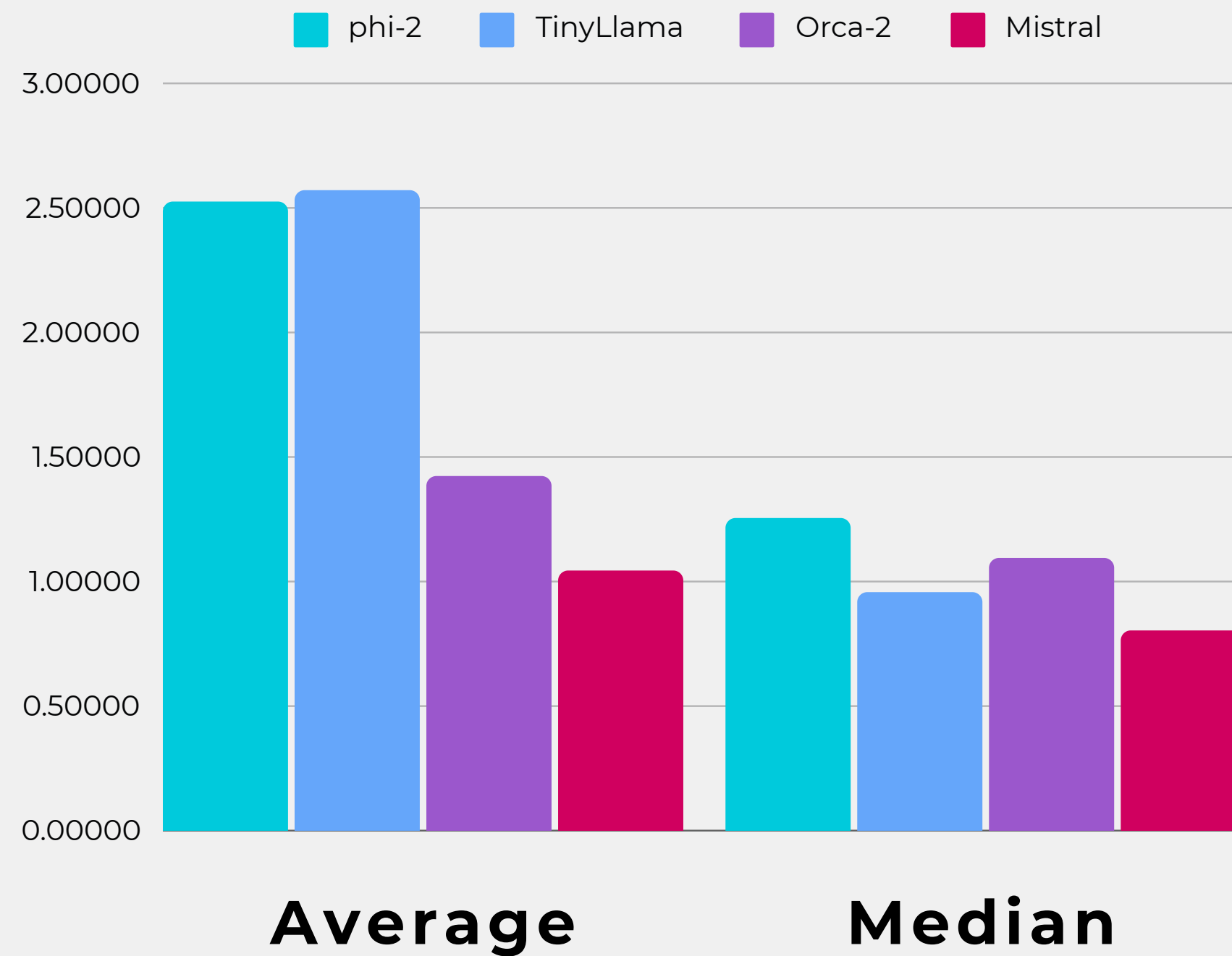


06. Accuracy e Edit Distance

Accuracy



Edit Distance



06. Perplexity

LLM	PARAMETRI	MEDIA PERPLEXITY	VARIANZA PERPLEXITY
phi-2	2.7B	460	2,0699 E+4
orca-2	13B	205	1,6395 E+5
Mistral	7B	208	2,9106 E+5
TinyLlama	1.1B	12494	8,7219 E+8

06. Dati qualitativi

LLM	PARAMETRI	EPOCHE	RISULTATO QUALITATIVO	TEMPO RISPOSTA
phi-2	2.7B	3	✓	5 min
orca-2	13B	3	✓	14 min
Mistral	7B	10	✓	23 min
TinyLlama	1.1B	20		9 min
OpenHermes	7B	10	✓	2h 28 min

● 07. Conclusioni

Orca-2 e phi-2 hanno ottenuto prestazioni del **50%** migliori sugli altri LLM di taglia media e di **60 volte** migliori sul meno performante TinyLlama, ma **phi-2**:

- è migliore nel test qualitativo
- è più piccolo (risparmio del 70%)
- ha ottenuto prestazioni allo stato dell'arte pari di un LLM di taglia **medio-alta** (Orca-2)

SMALL BUT MIGHTY

Phi-2

2.7 billion parameters

Common sense, language understanding,
and logical reasoning



GRAZIE DELL'ATTENZIONE

