



UNIVERSITÁ DI VERONA

# USO DEGLI LLM PER ANALIZZARE E PREDIRRE TRAIETTORIE TURISTICHE

Mattioli Simone

# TABELLA DEI CONTENUTI

---

01 Introduzione Al Problema

02 Obiettivi

03 Metodologia

04 Strategie Di Promting

05 Ambiente Di Sviluppo

06 Modelli Considerati

07 Risultati Ottenuti

08 Conclusione

# INTRODUZIONE AL PROBLEMA

01



# PREDIRRE

---



## SFIDE ATTUALI

- Pattern Complessi
- Influenze Multiple

## OPPORTUNITÀ

- LLM e Le Loro Capacità Emergenti
- Capacità Di Calcolo

# OBIETTIVI

---

02

# COSA SI VUOLE RAGGIUNGERE?

---



## Performance Temporali

Deve Essere Utilizzabile



## FrameWork Indipendente dall' LLM

LLM Intercambiabile



## Precisione Su Dati Reali

Test sulle VeronaCard

# METODOLOGIA

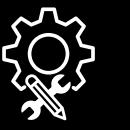
---

03



## RICERCA PRELIMINARE

FrameWork Pre-Esistenti



## SVILUPPO

Nuove funzionalità



## VALIDAZIONE

Feature Implementata

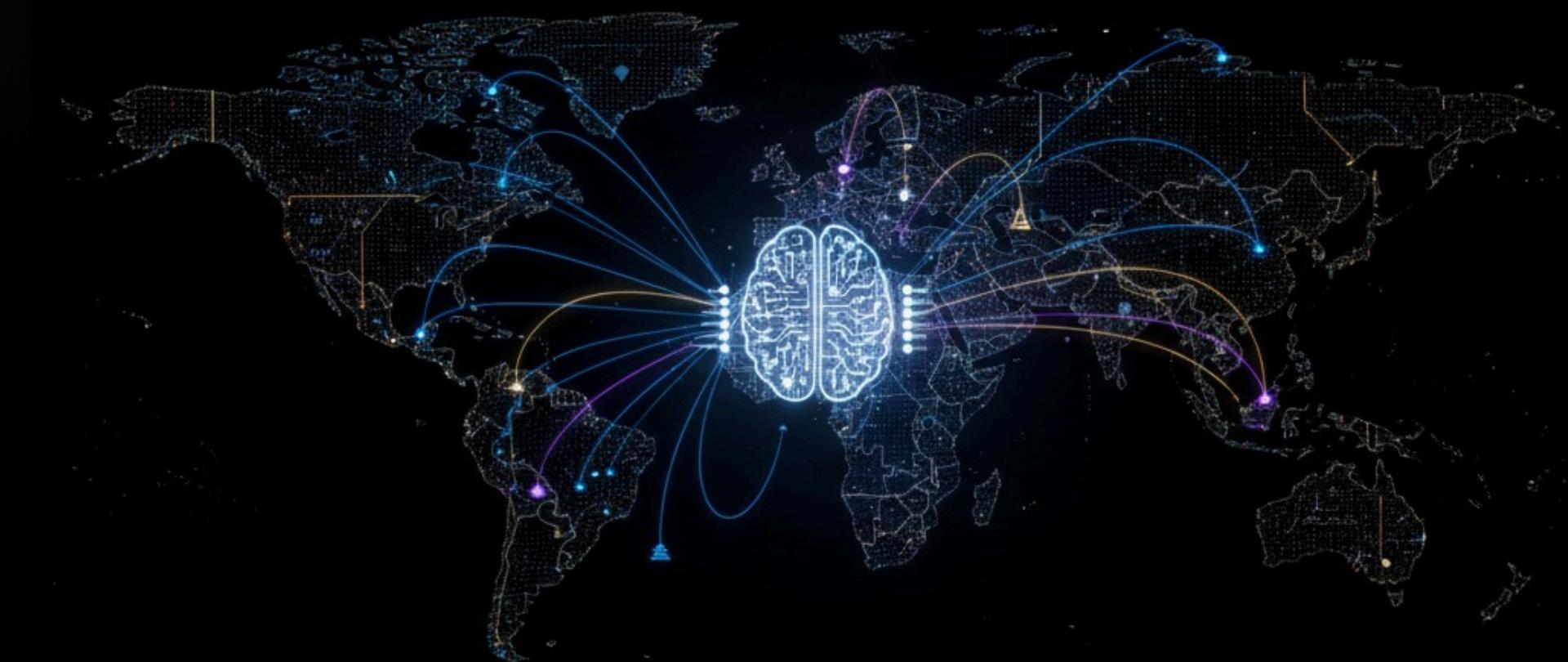


## ANALISI PRESTAZIONI

Con Metriche Specifiche

# PIANO STRATEGICO

01



# STEP PRATICI EFFETTUATI

—  
02



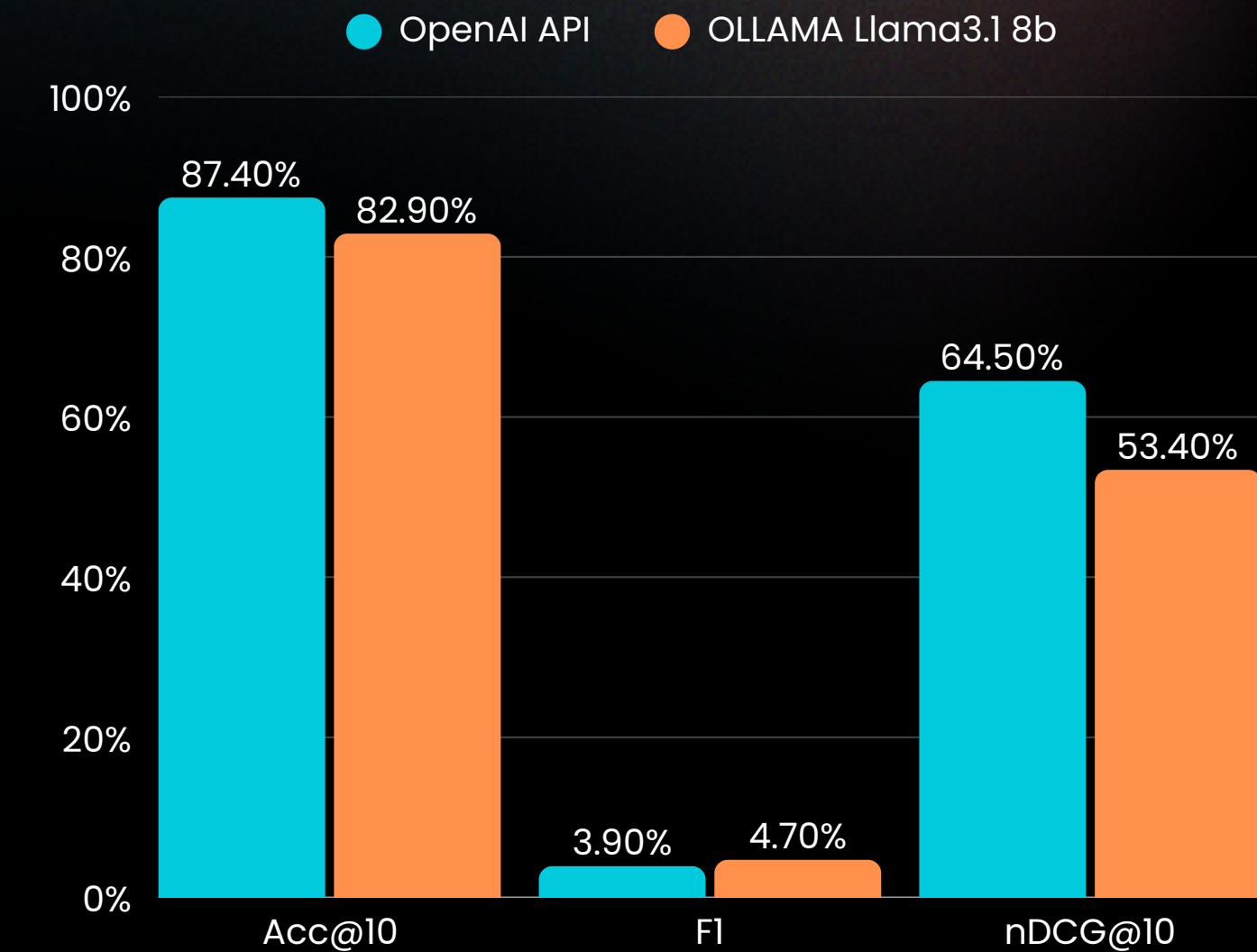
## OpenAI API KEY

- Abbonamento
- No Privacy
- SOLO Modelli Di OpenAI



## OLLAMA Llama3.1 8b

- Privacy
- Totalmente Gratuito
- Sicurezza dei dati
- TUTTI i Modelli Open Source



# DATASET VERONACARD

---

03

## ANAGRAFICA DEI POI ( Point Of Interest )

- 22 Attrazioni Turistiche Di Verona
- Coordinate Geografiche
- Nomi Completi
- Nomi Abbreviati

## LOG TRANSAZIONALI DEGLI ACCESSI

- Divisi Per Anni ( 2014 → 2023 )
- TimeStamp Ingresso
- POI Visitato
- ID VeronaCard



# STRATEGIE DI PROMTING

---

04

# APPROCCIO INCREMENTALE

---

BASE

NOME

- Promt Semplice

GEOM

NOME  
COORDINATE

- Promt Medio

TIME

NOME  
COORDINATE  
TIMESTAMP

- Promt Complesso

# AMBIENTE DI SVILUPPO

05

# HPC LEONARDO

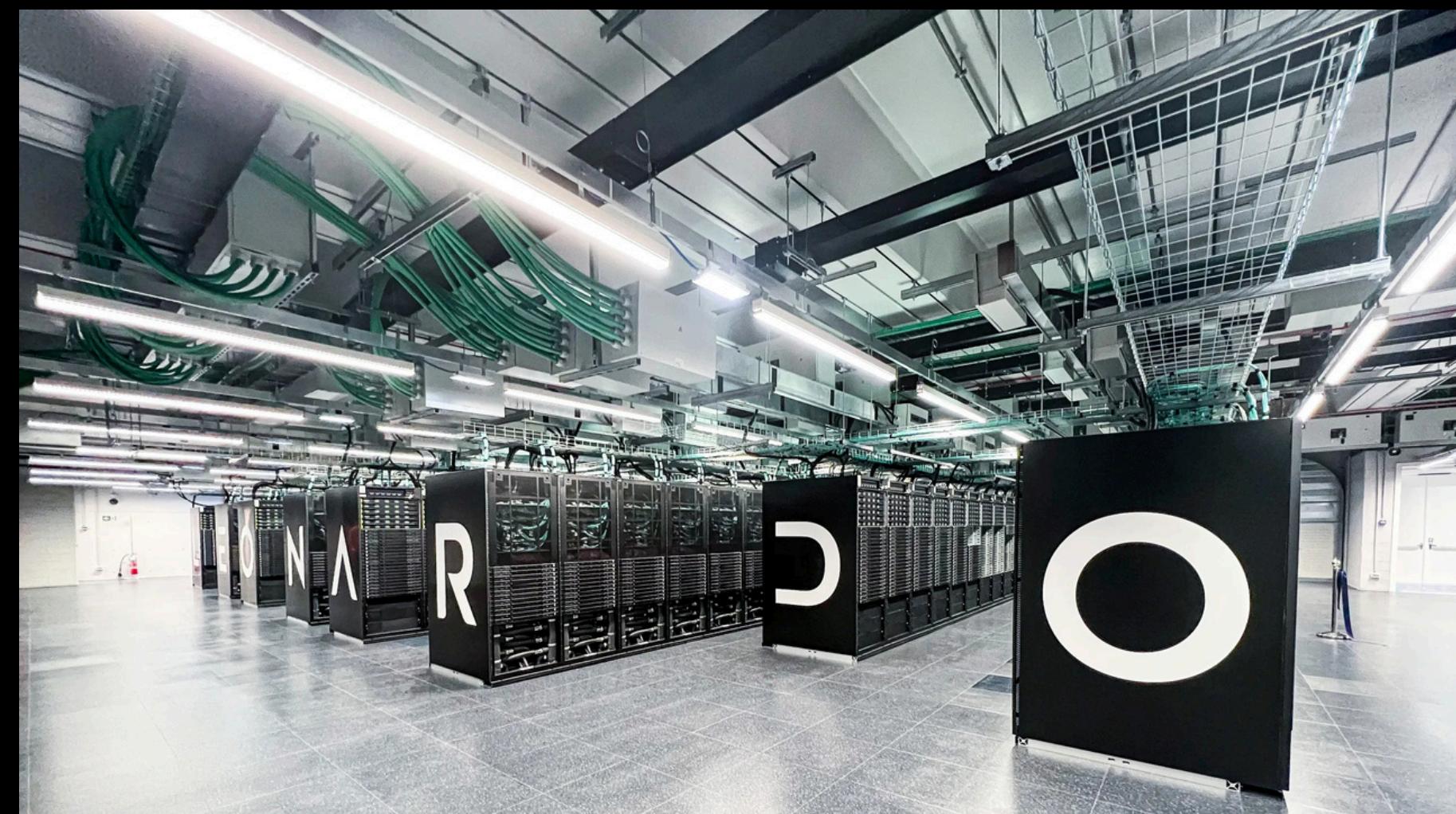
BOOSTER MODULE

## Architettura

- 3456 NODI
- 512GB Di RAM DDR4 Per Nodo
- 4x NVIDIA custom Ampere A100 Per Nodo

## Configurazione Utilizzata

- Ollama Multi-Istanza su 4 GPU
- 12 Richieste Concorrenti ( 3 x GPU )
- Checkpoint system





## DOCUMENTAZIONE

Studio per l'utilizzo corretto



## SVILUPPO

Nuove funzionalità



## DEVELOPMENT

Attesa allocazione risorse



## ESECUZIONE

Su Leonardo

# SPEED UP OTTENUTO

Grazie all'HPC Leonardo

## PRIMA

- 1 File Al Giorno
- 2 Secondi Per Card
- Settimane Di Esecuzione

## DOPO

- Da 7 a 12 Cards Al Secondo
- Meno Di 24h Per Esecuzione

# MODELLI CONSIDERATI

---

06

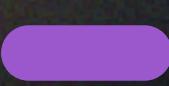
# LLAMA 3.1

---

7B

- Molto Efficiente
- SOTA Meta ( baseline )
- Utilizzato per sviluppo e testing

# QWEN 2.5



**7B**

- Ottima Comprensione Contestuale
- Eseguibile su Hardware Consumer



**14B**

- Impatto Scaling

# MISTRAL

---

## MISTRAL 7B

- Molto Efficiente su singola GPU
- Test Tra Modelli Piccoli

## MIXTRAL 8x7B

- Architettura MoE
- Test Tra Modelli Medio/Piccoli

# DEEPSEEK

---

## CODER 33B

- Modello Specializzato
- Alternativa Tra Modelli Medio/Piccoli

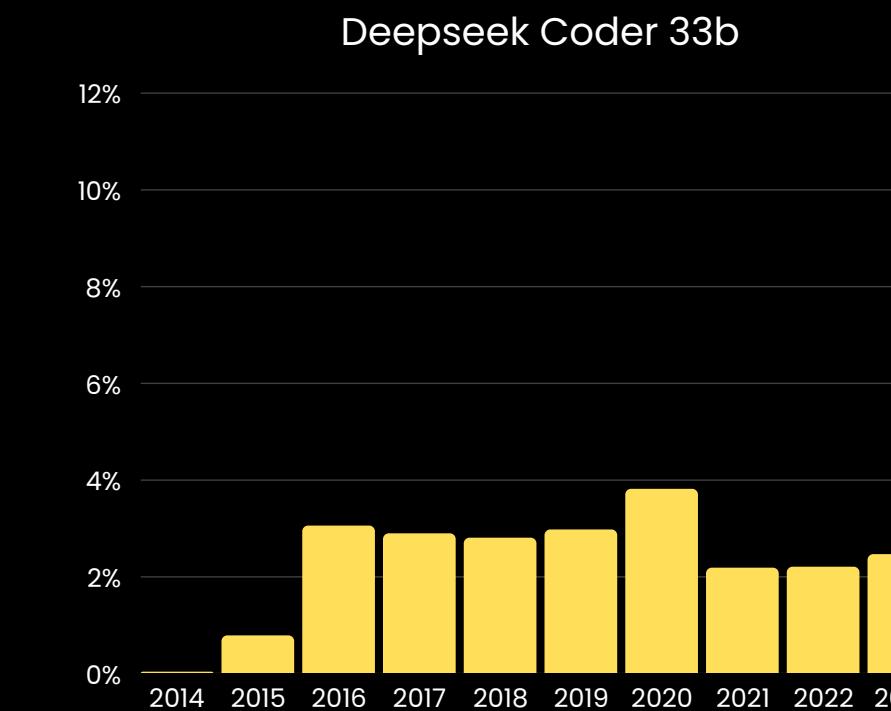
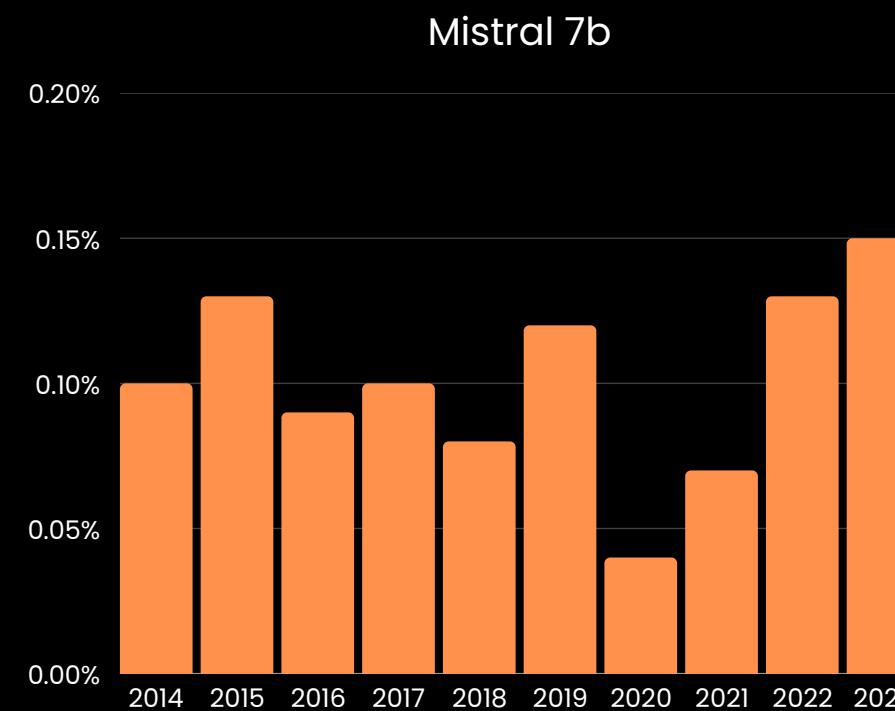
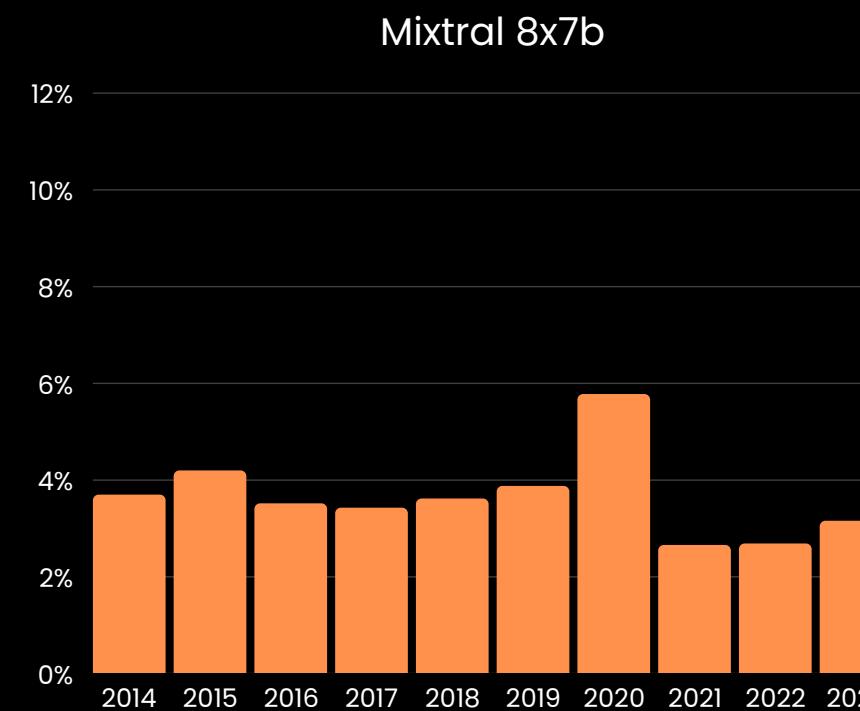
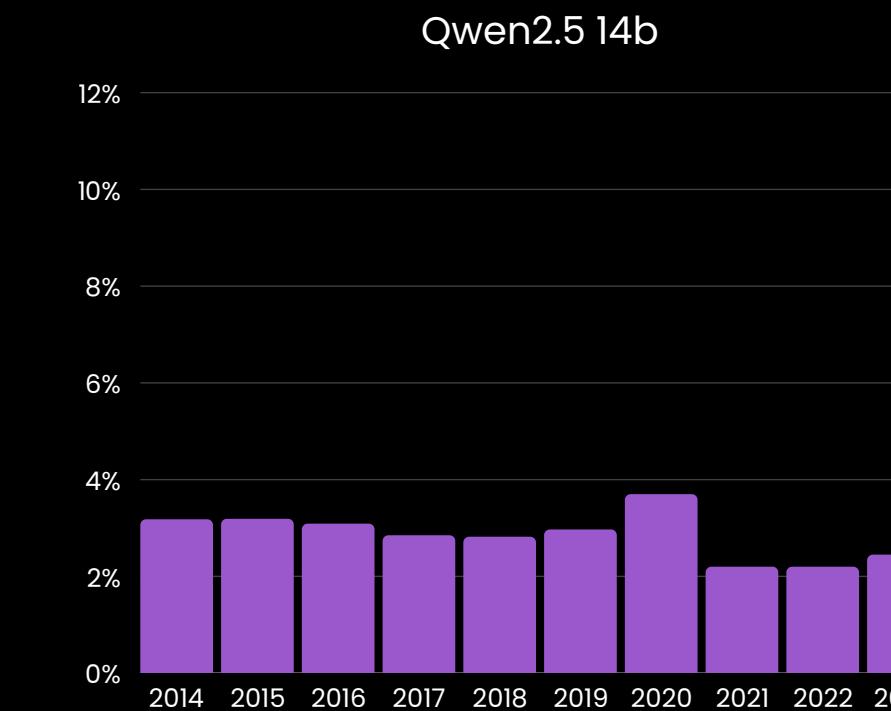
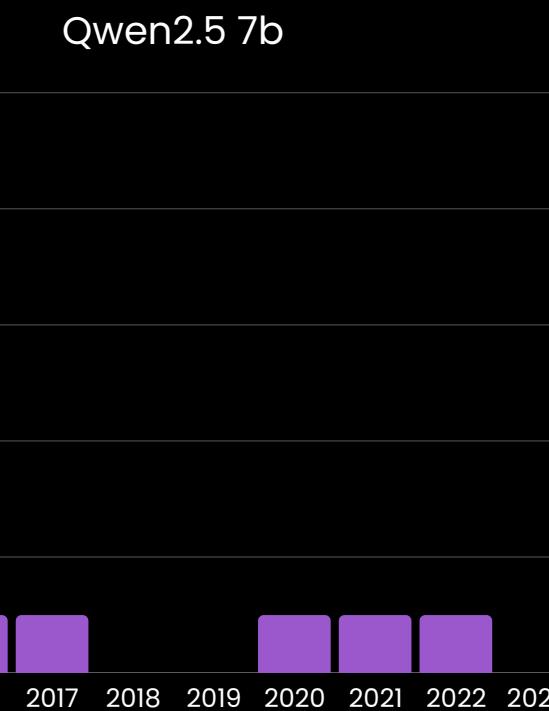
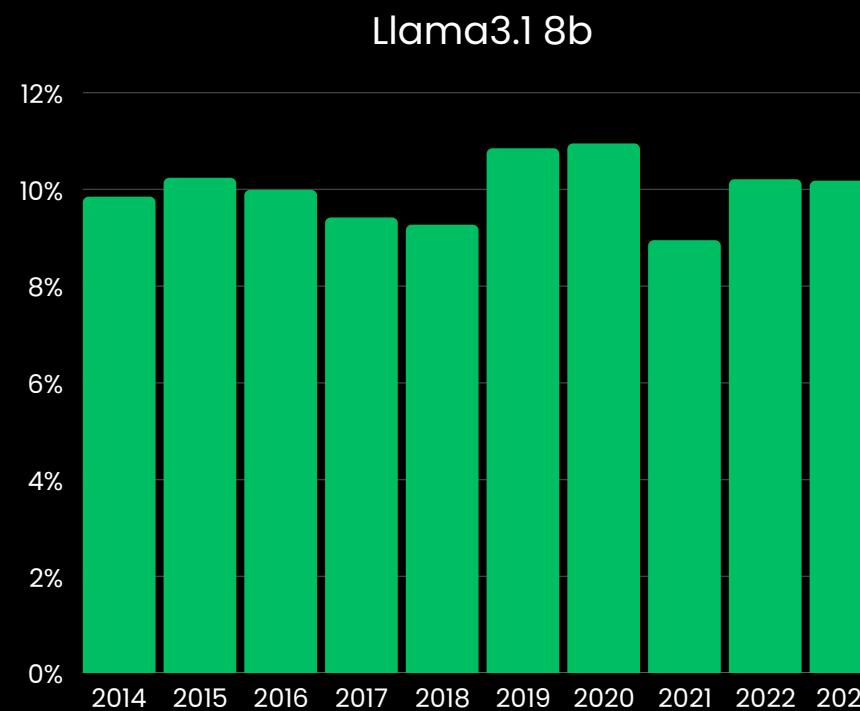
# RISULTATI OTTENUTI

---

07

# STRATEGIA BASE - PENULTIMO

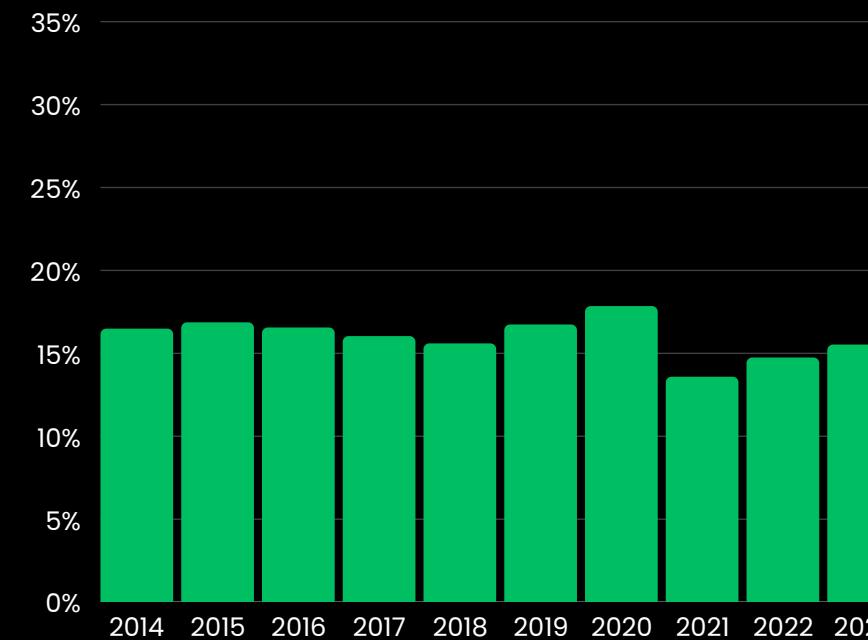
TOP 1 ACCURACY



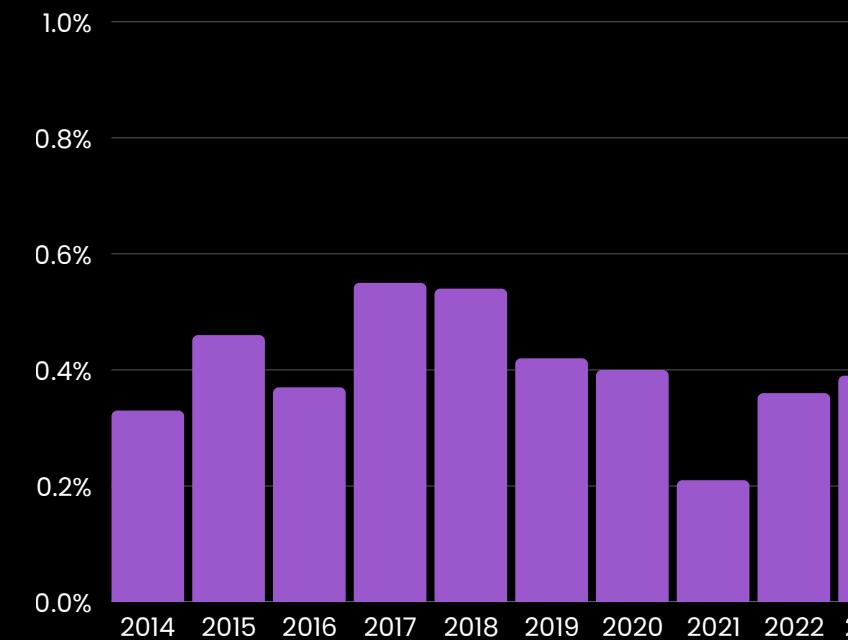
# STRATEGIA BASE - PENULTIMO

5 HIT RATE

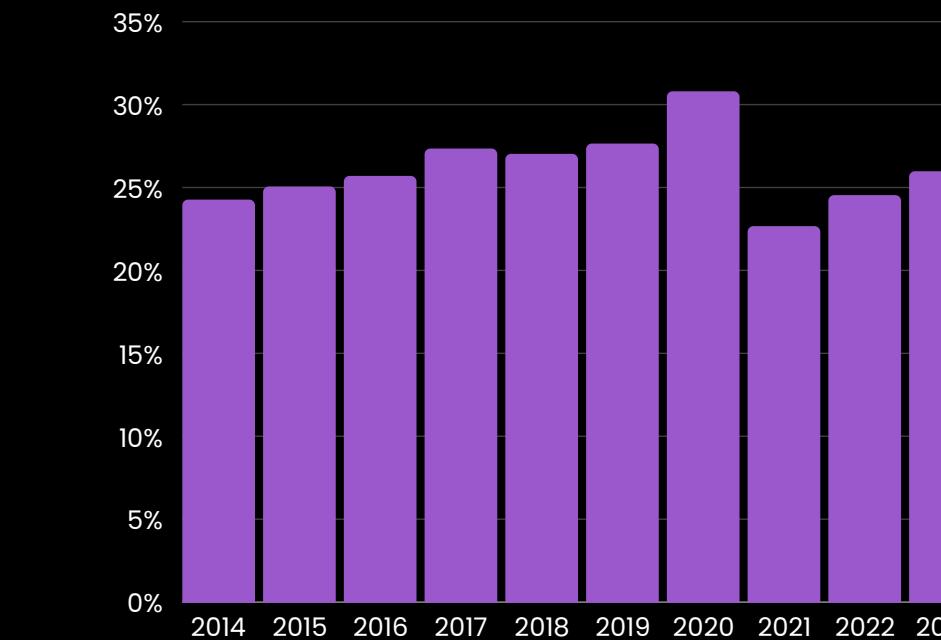
Llama3.1 8b



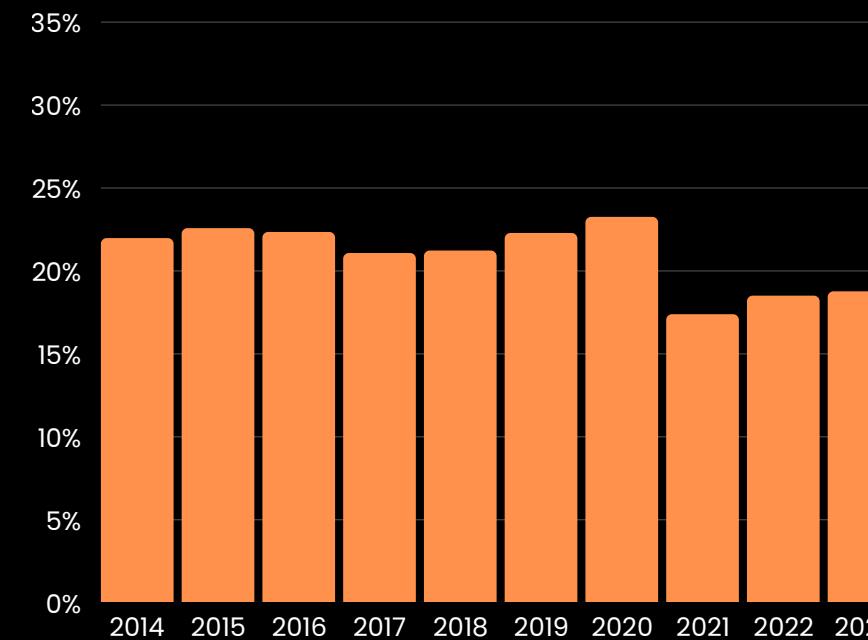
Qwen2.5 7b



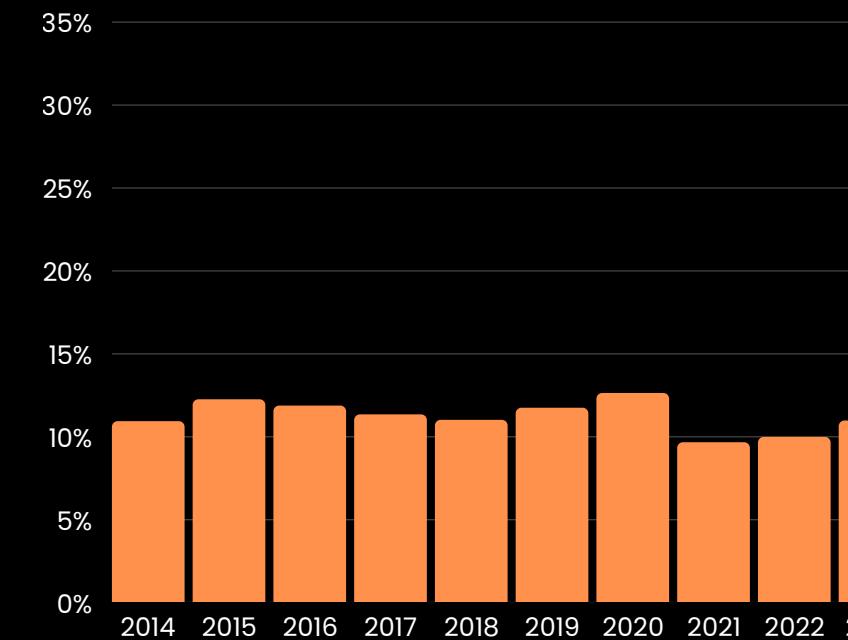
Qwen2.5 14b



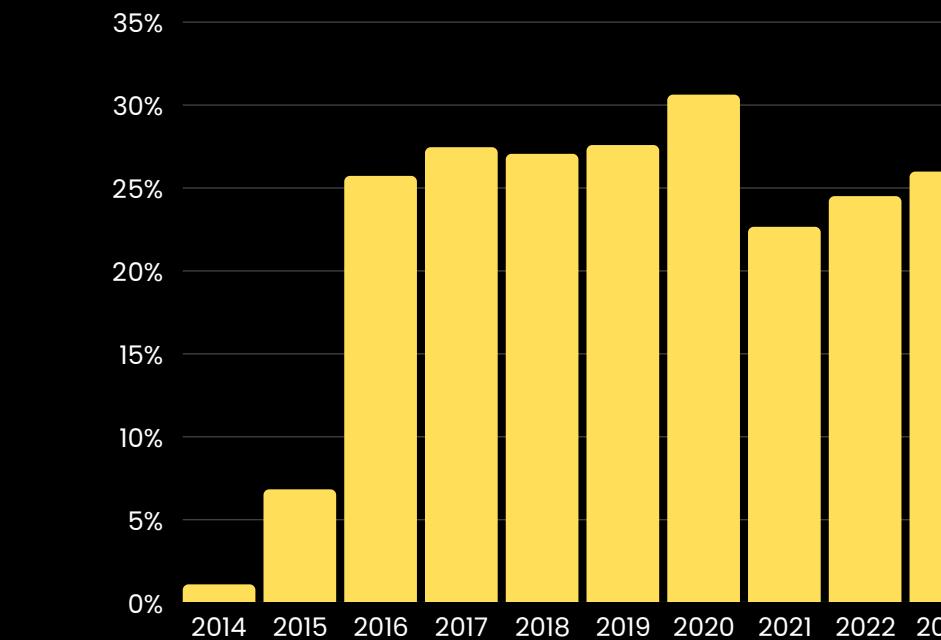
Mixtral 8x7b



Mistral 7b

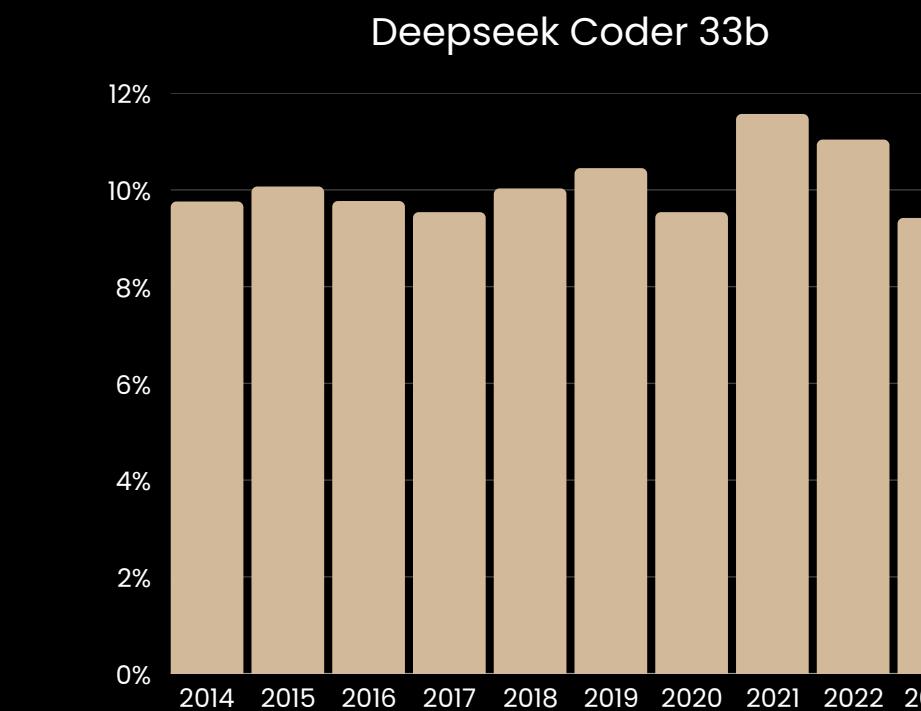
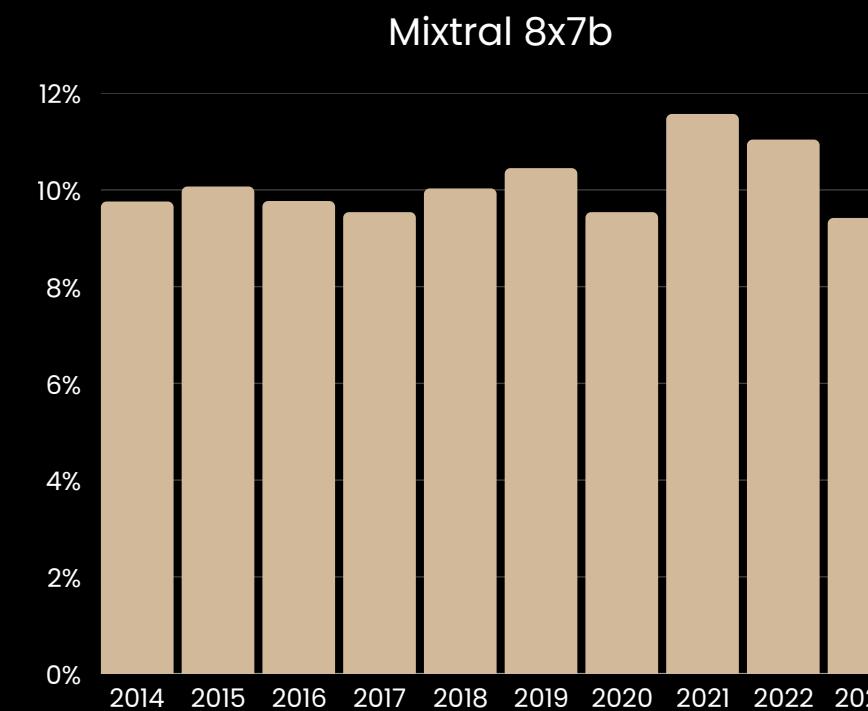
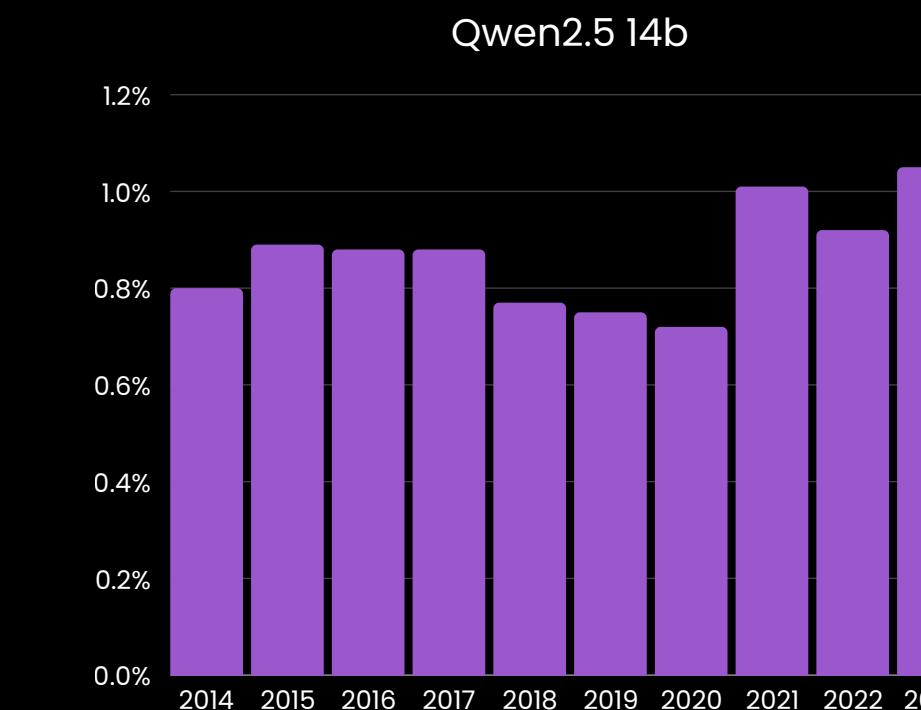
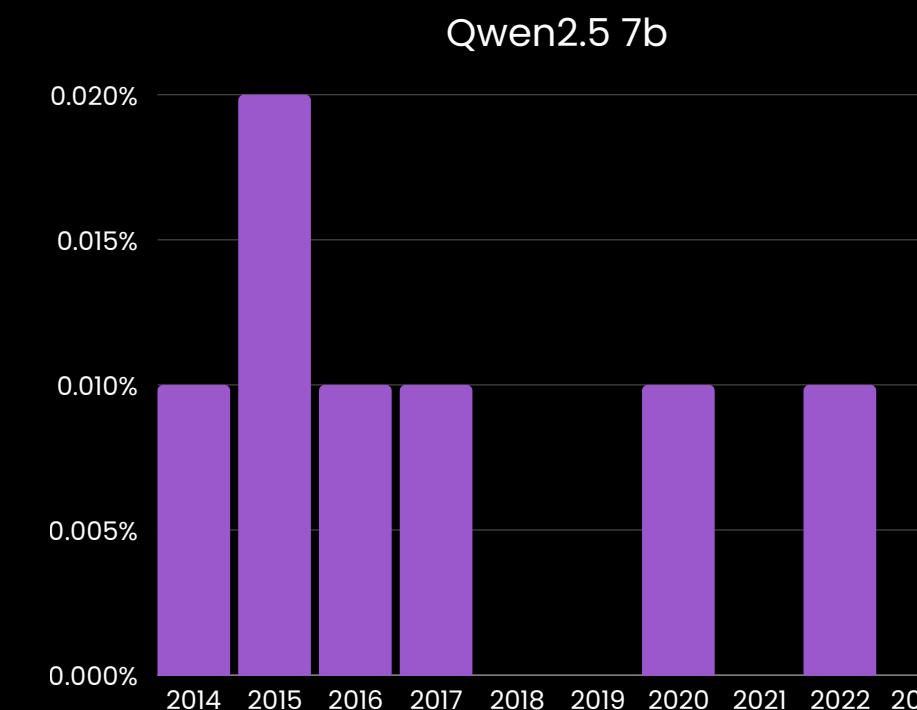


Deepseek Coder 33b



# STRATEGIA BASE - METÀ

TOP 1 ACCURACY



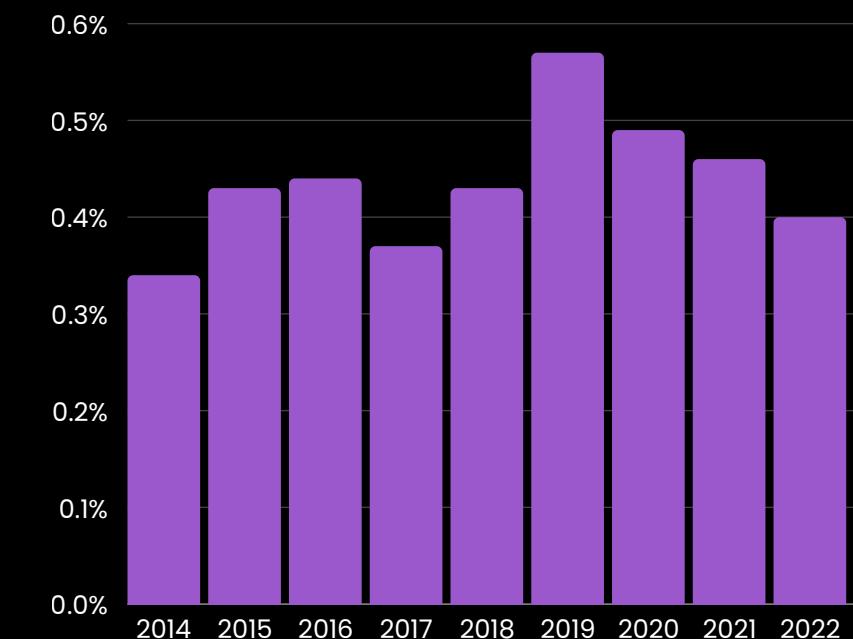
# STRATEGIA BASE - METÀ

5 HIT RATE

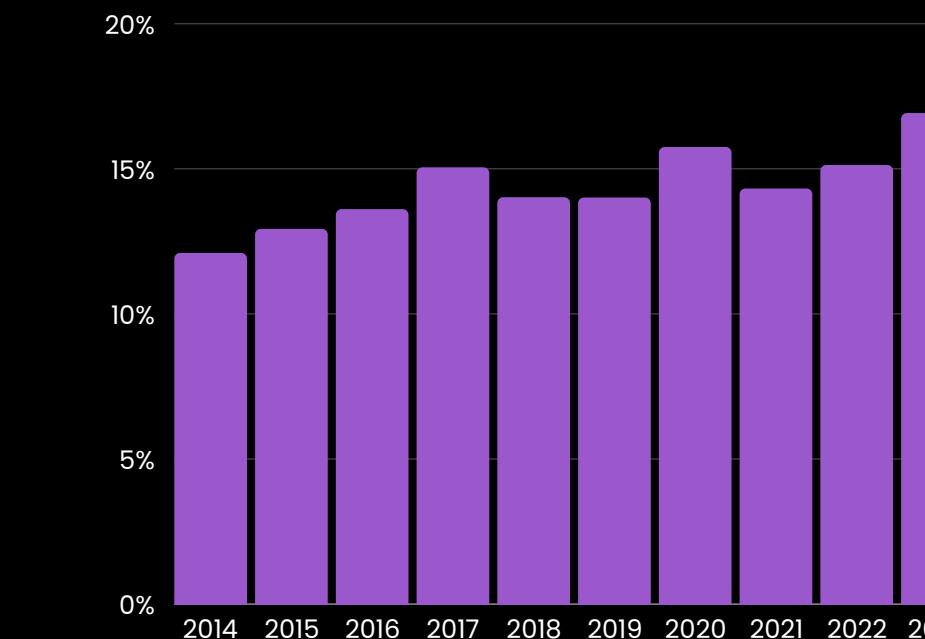
Llama3.1 8b



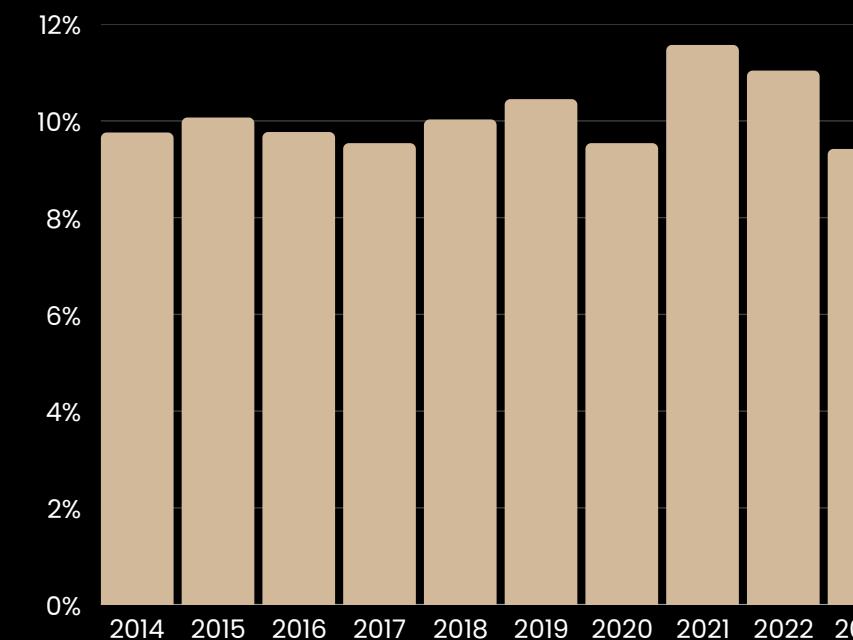
Qwen2.5 7b



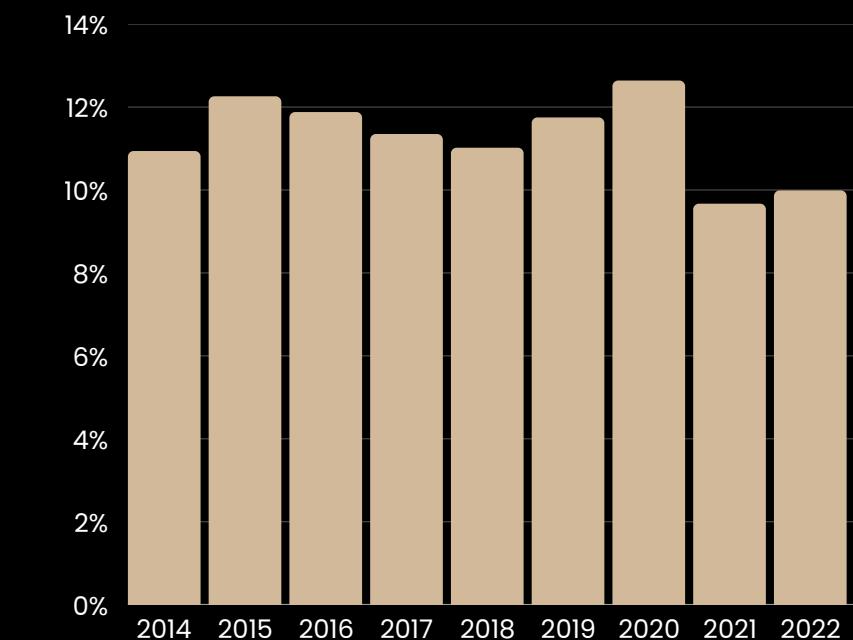
Qwen2.5 14b



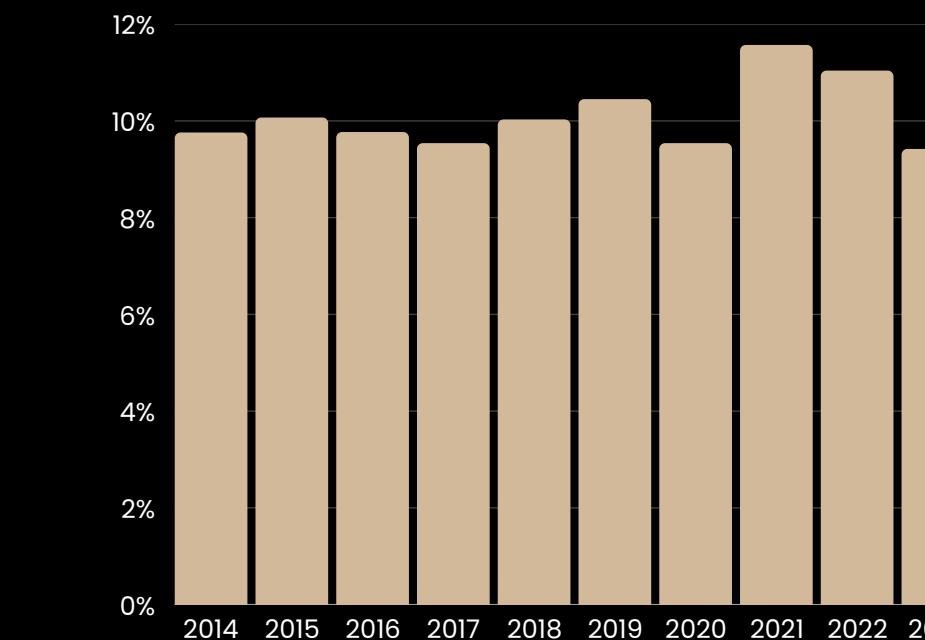
Mixtral 8x7b



Mistral 7b

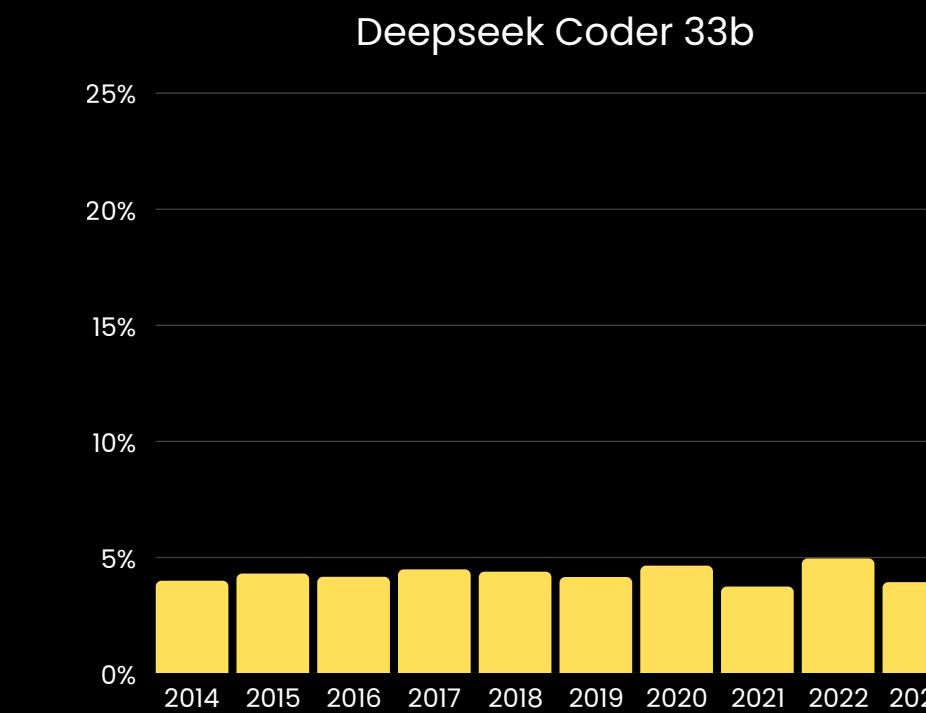
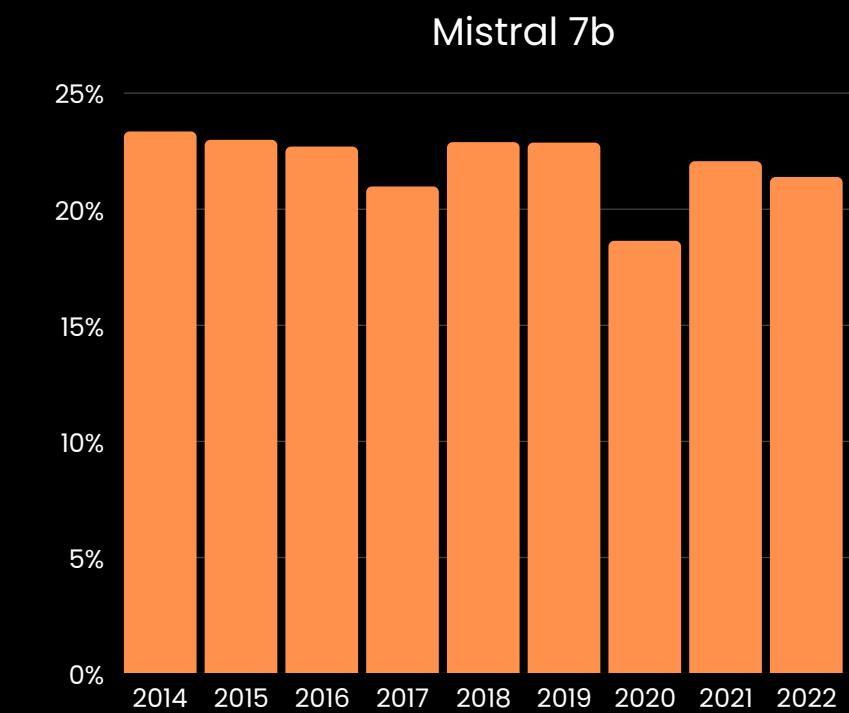
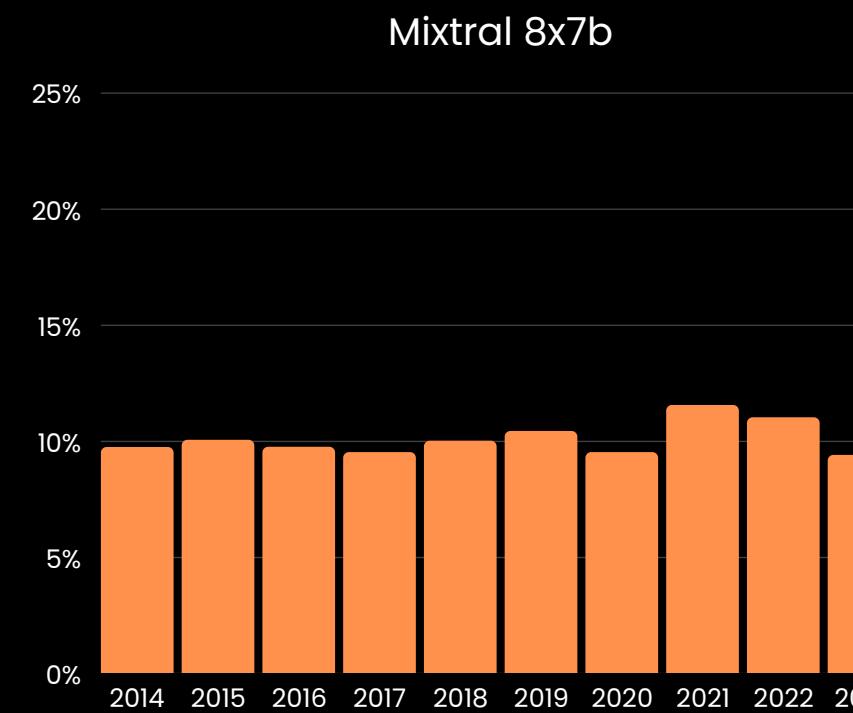
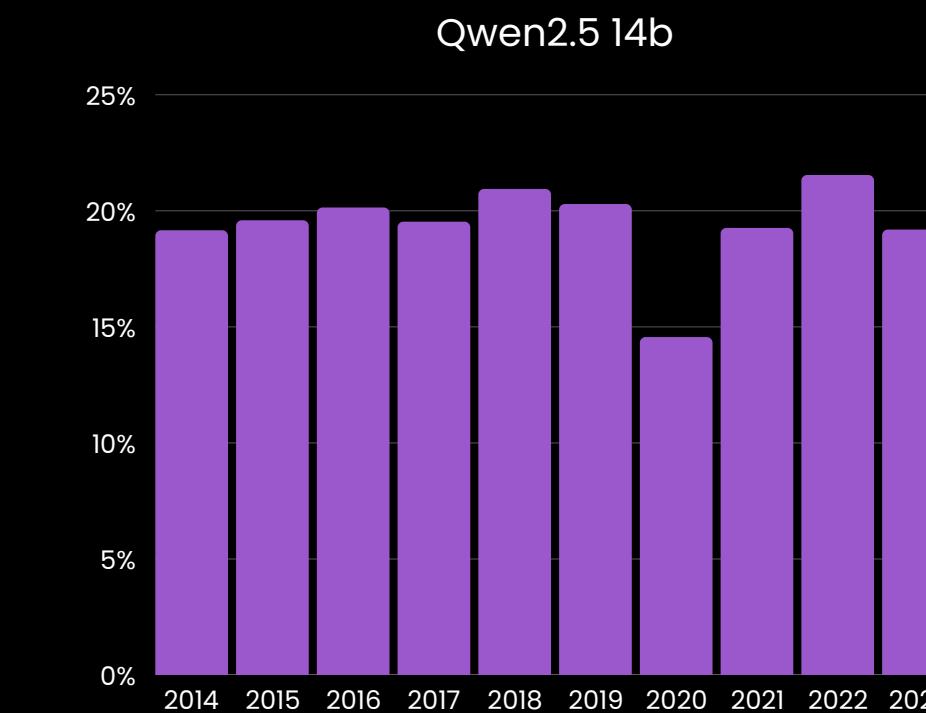
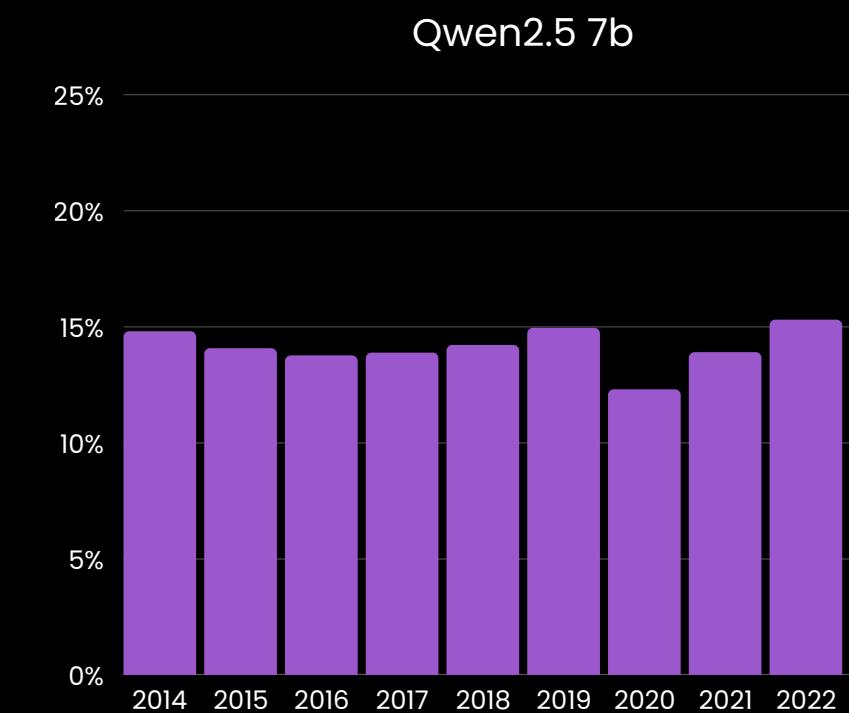
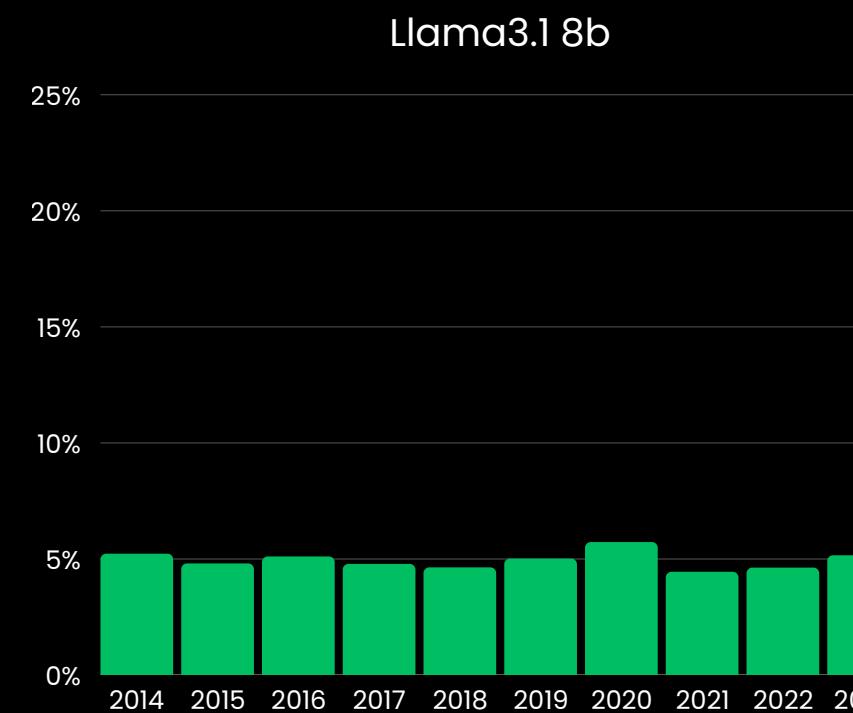


Deepseek Coder 33b



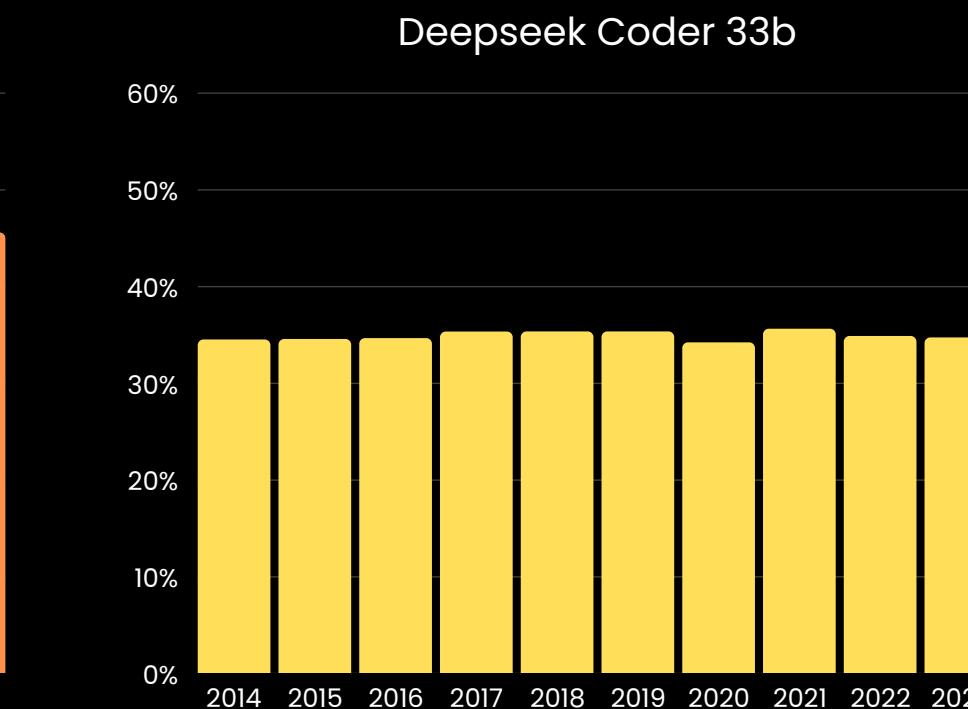
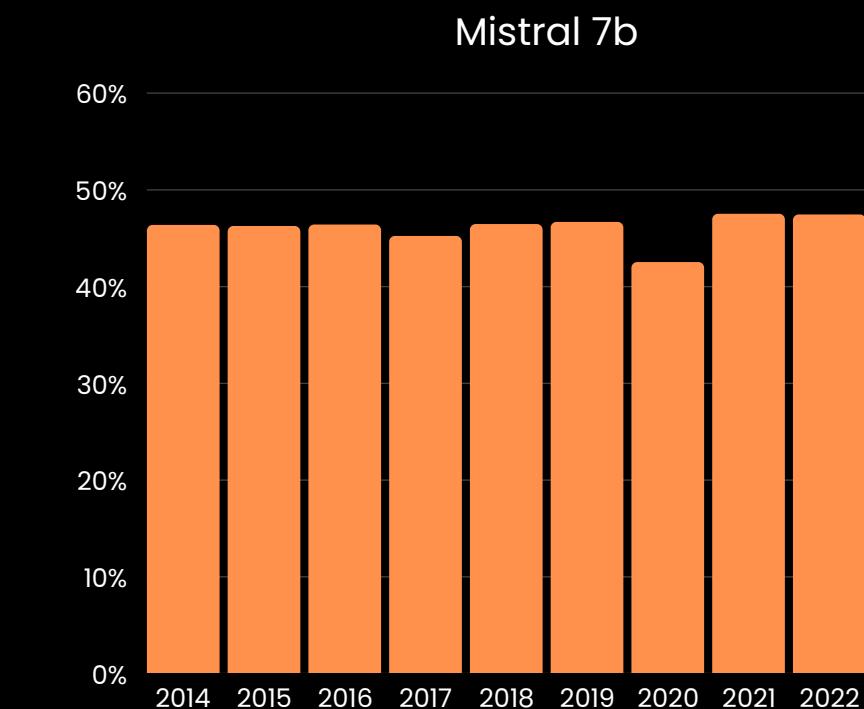
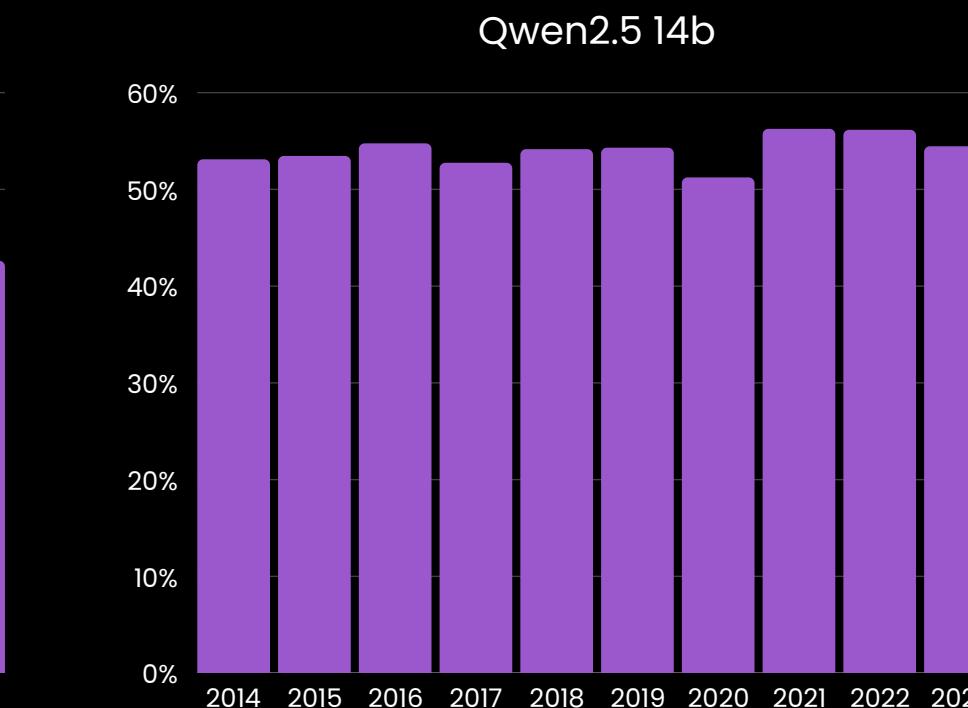
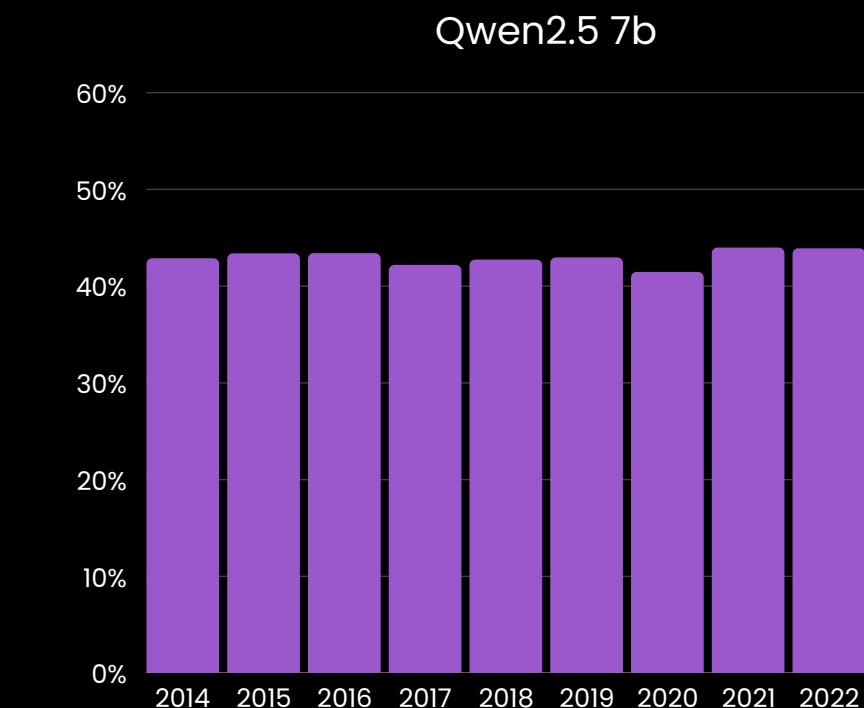
# STRATEGIA GEOM - PENULTIMO

TOP 1 ACCURACY



# STRATEGIA GEOM - PENULTIMO

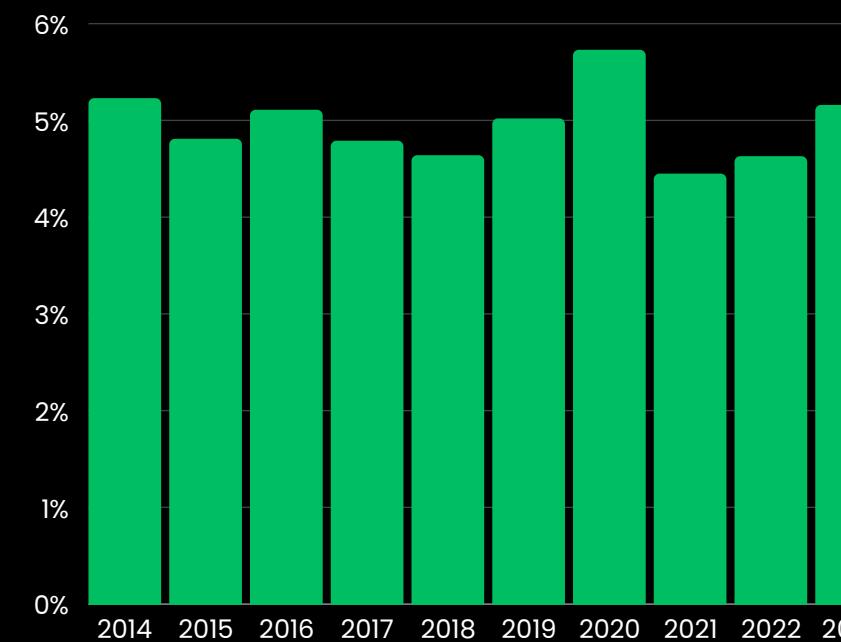
5 HIT RATE



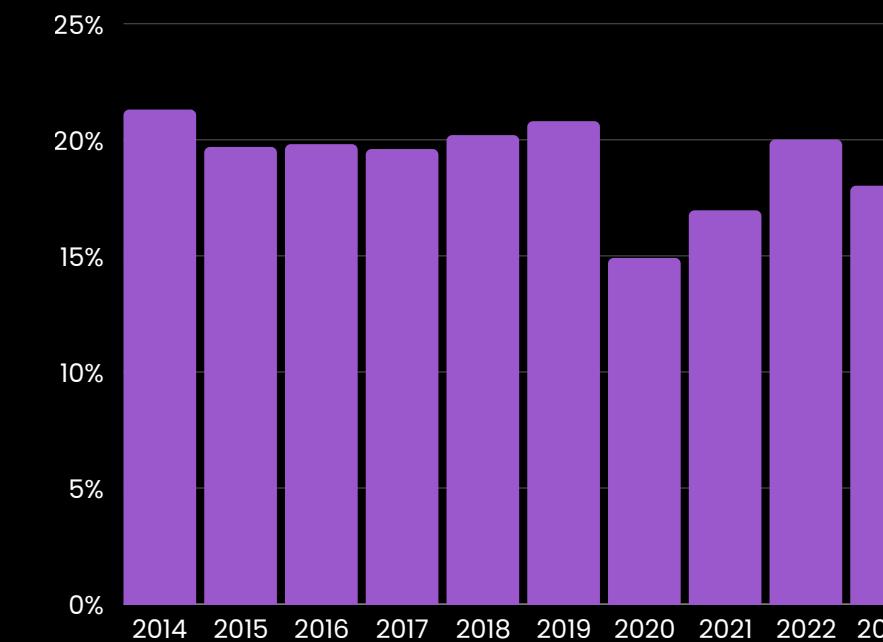
# STRATEGIA GEOM - METÀ

TOP 1 ACCURACY

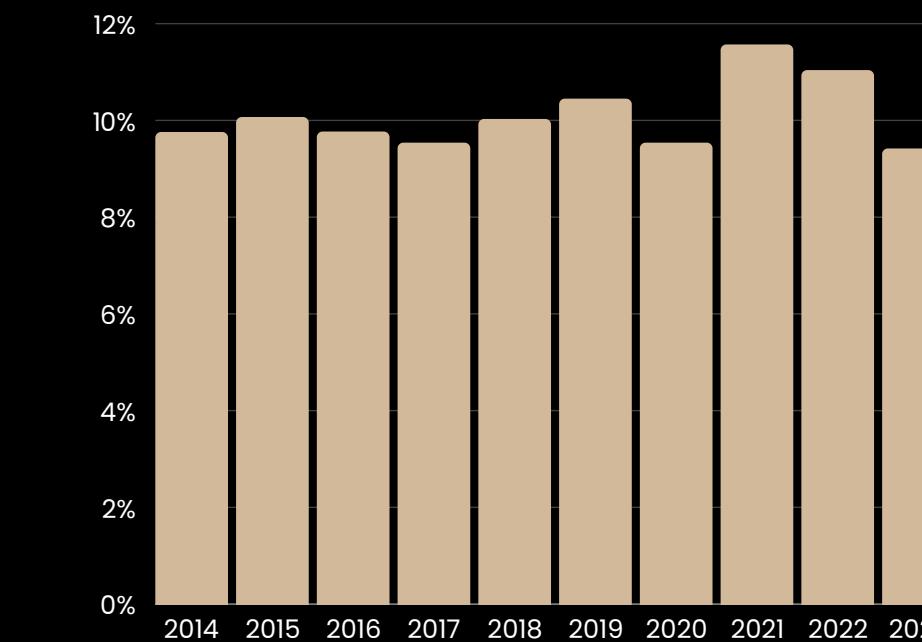
Llama3.1 8b



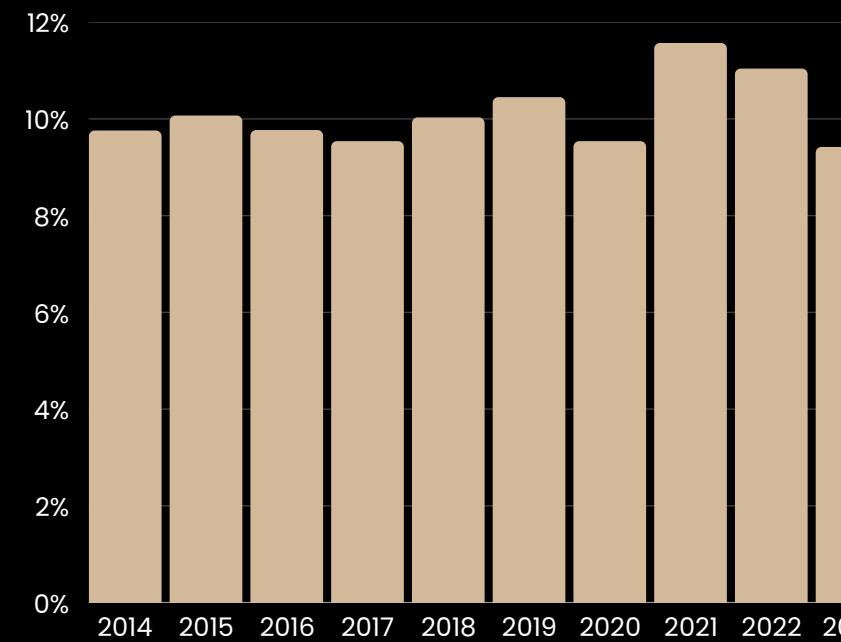
Qwen2.5 7b



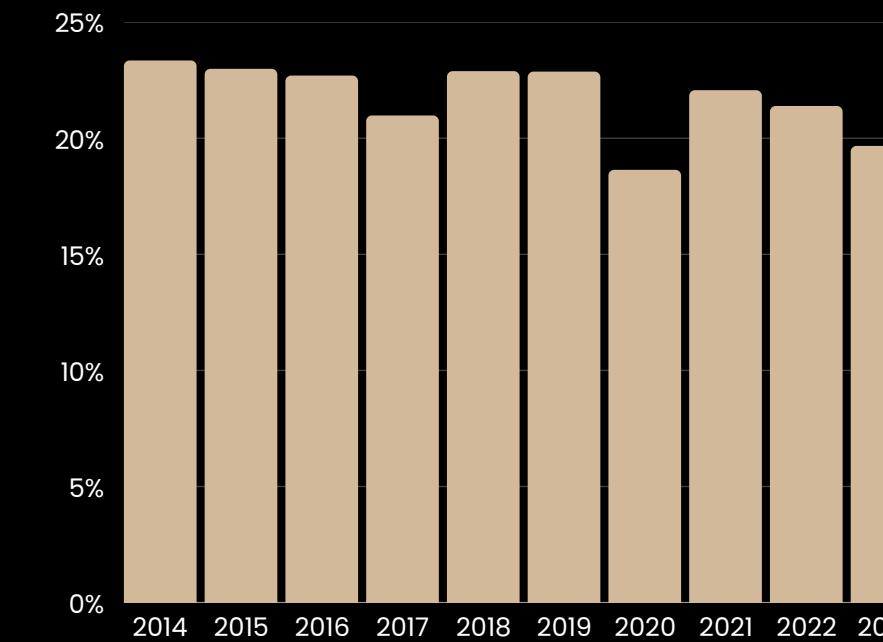
Qwen2.5 14b



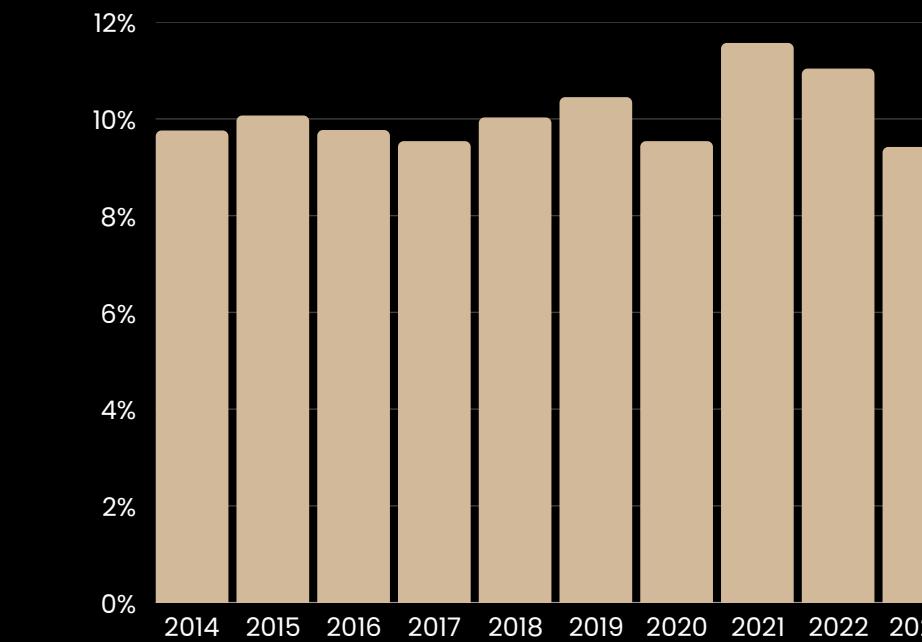
Mixtral 8x7b



Mistral 7b



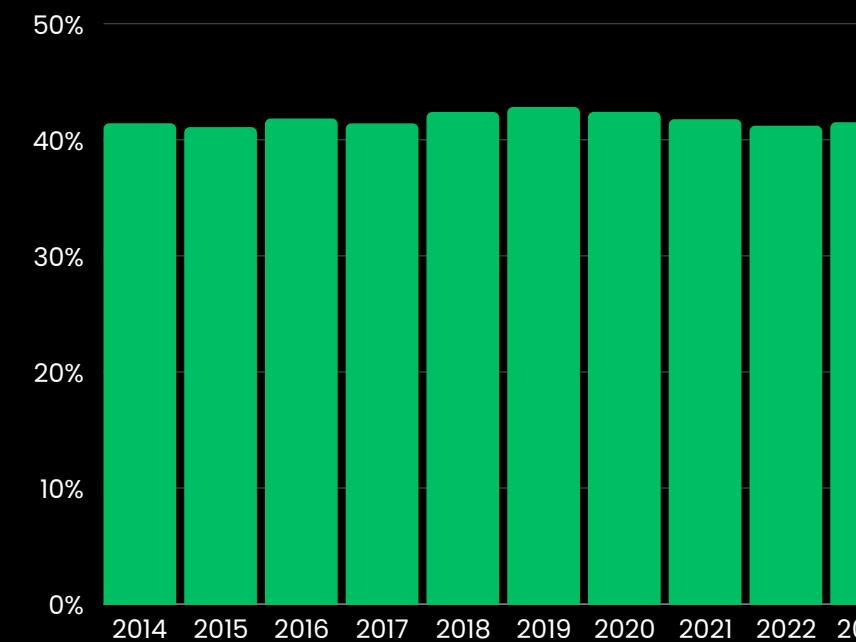
Deepseek Coder 33b



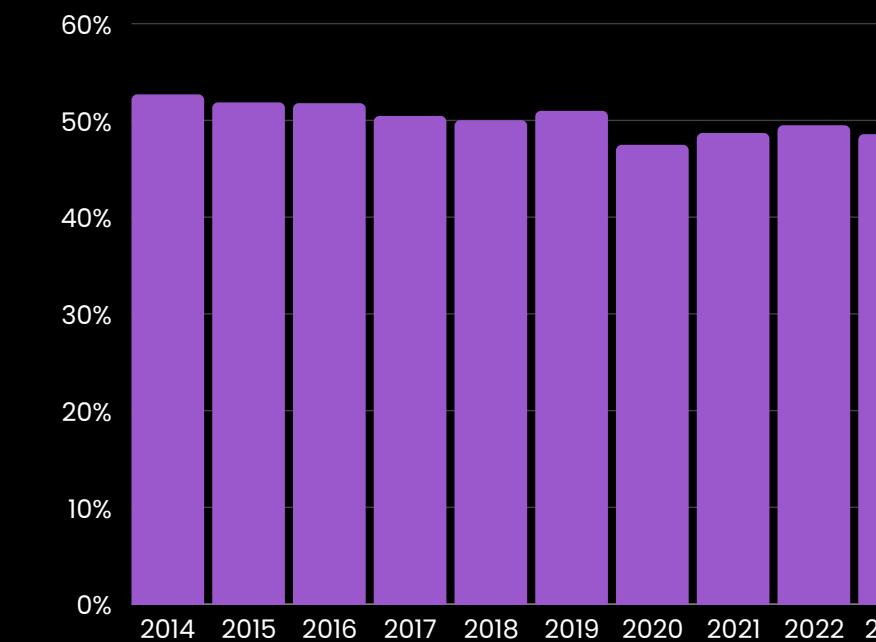
# STRATEGIA GEOM - METÀ

5 HIT RATE

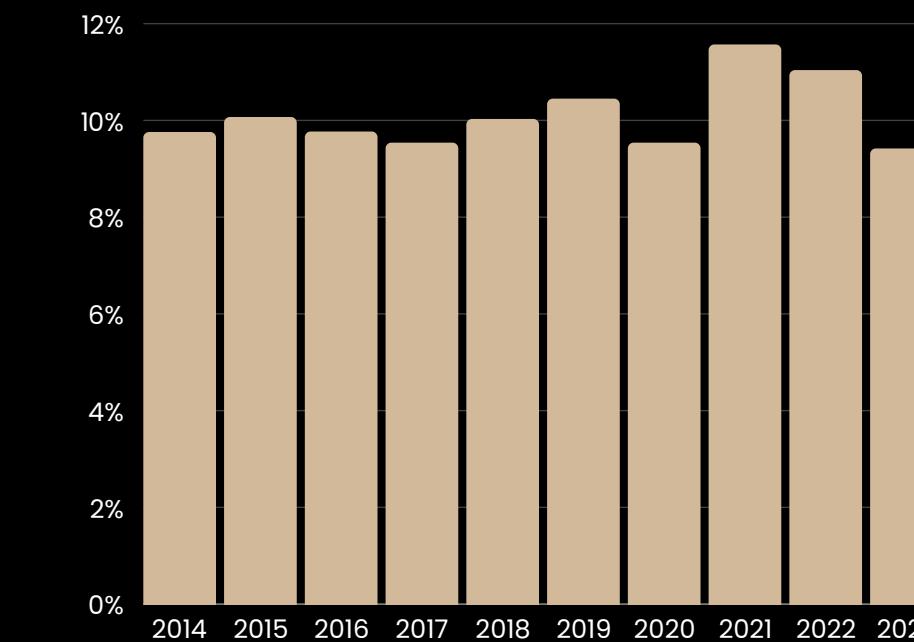
Llama3.1 8b



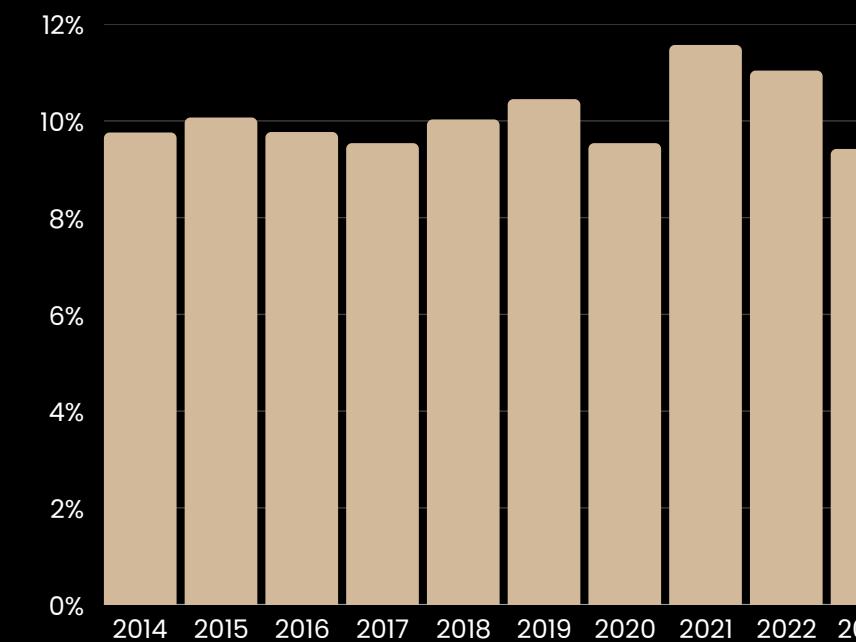
Qwen2.5 7b



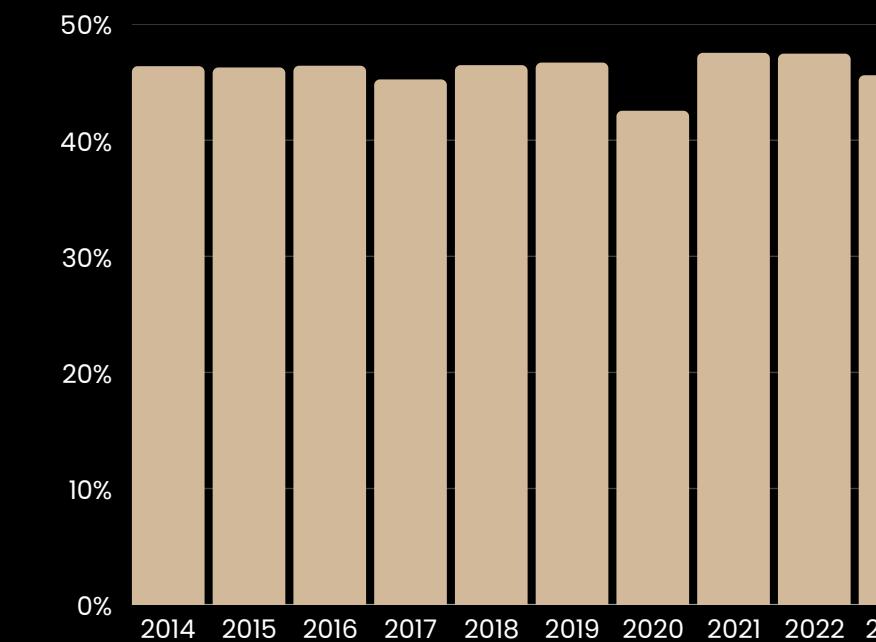
Qwen2.5 14b



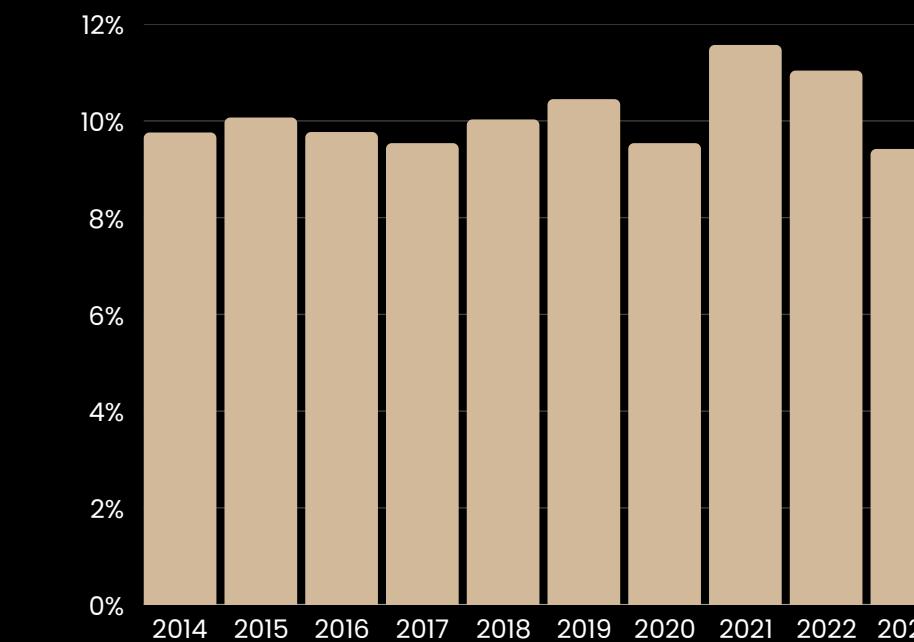
Mixtral 8x7b



Mistral 7b

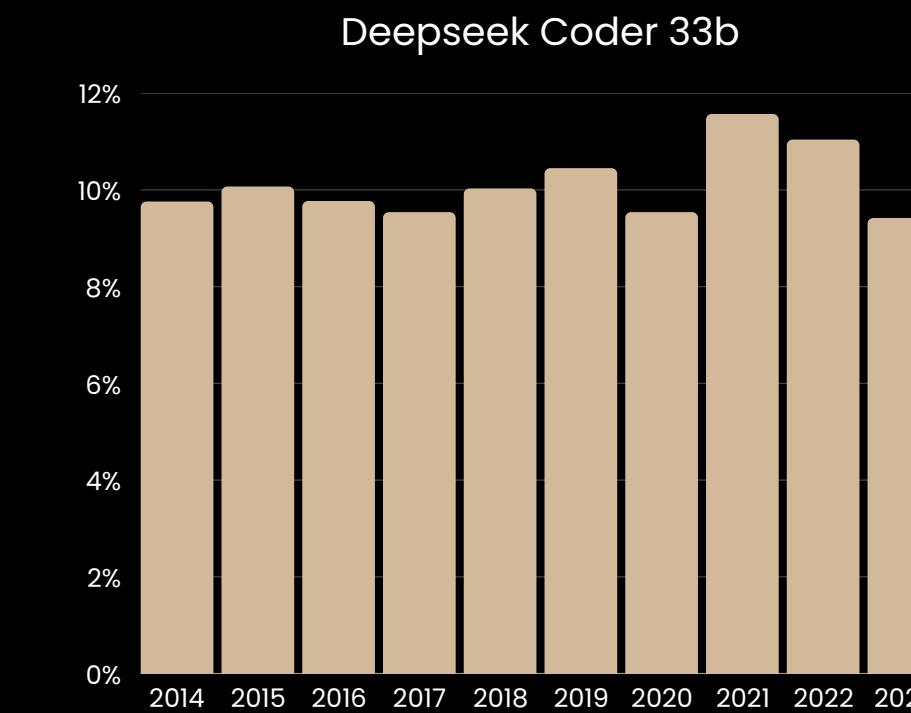
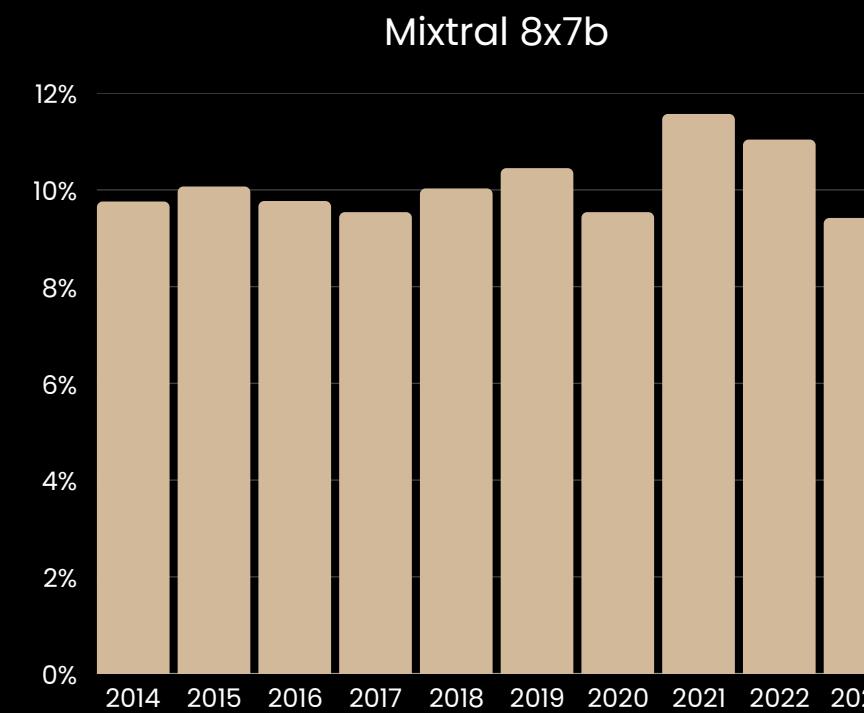
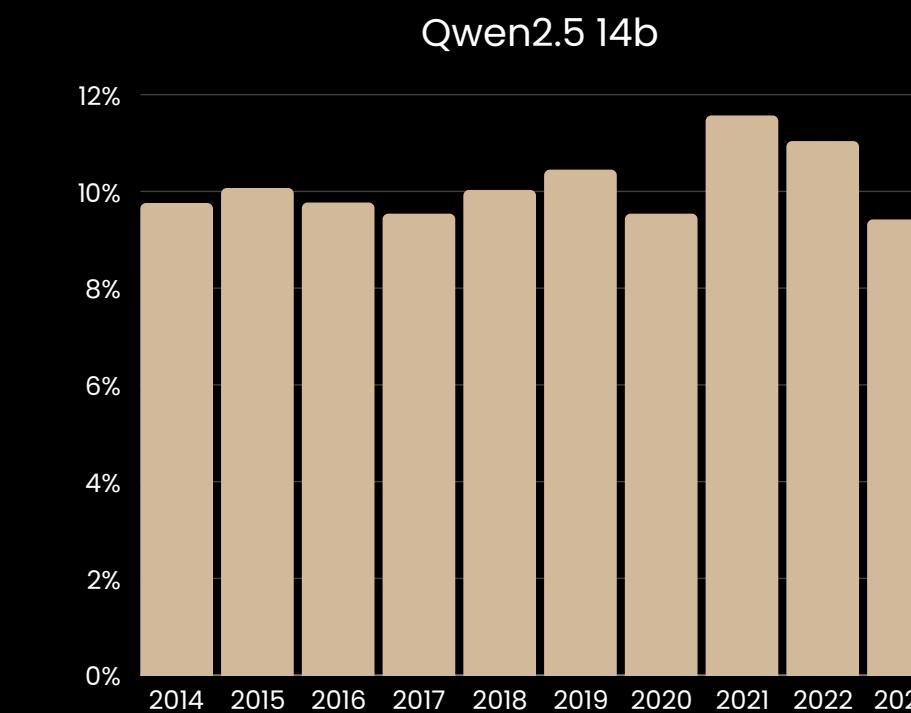
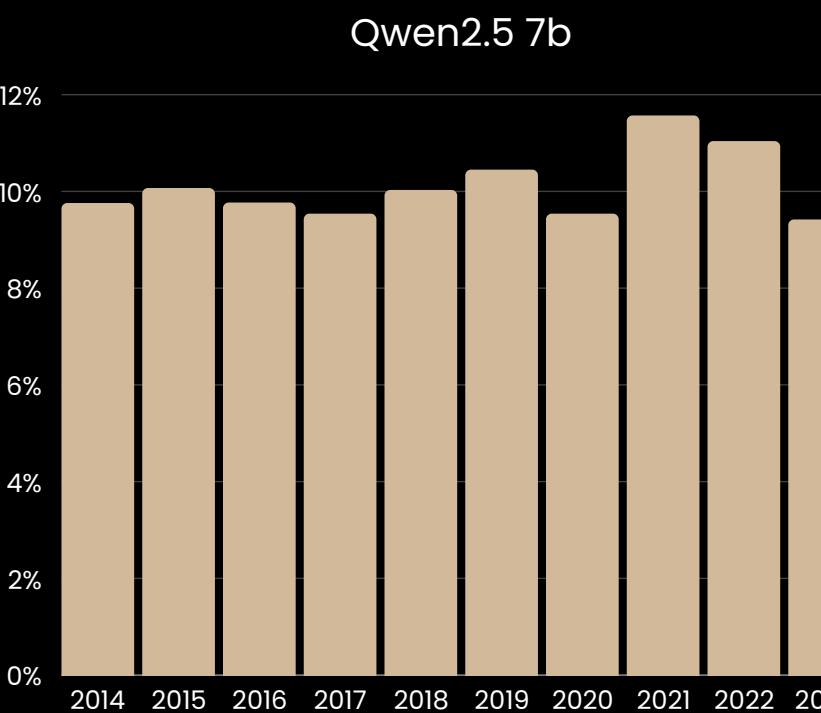
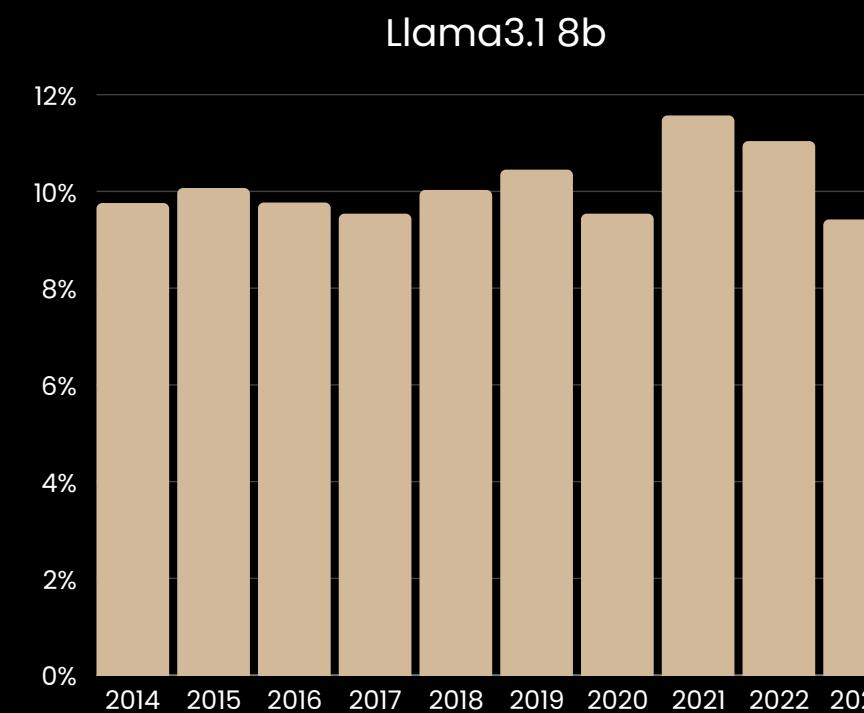


Deepseek Coder 33b



# STRATEGIA TIME – PENULTIMO

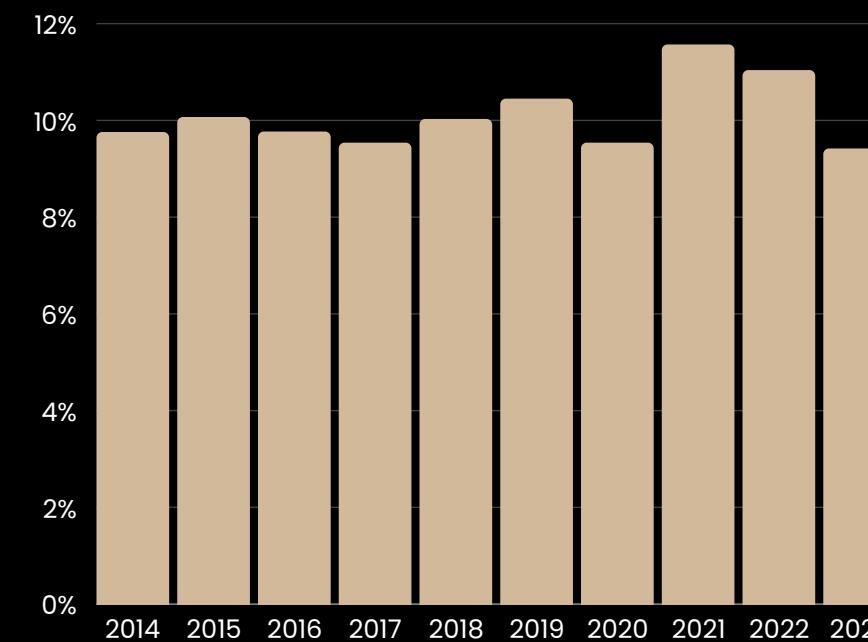
TOP 1 ACCURACY



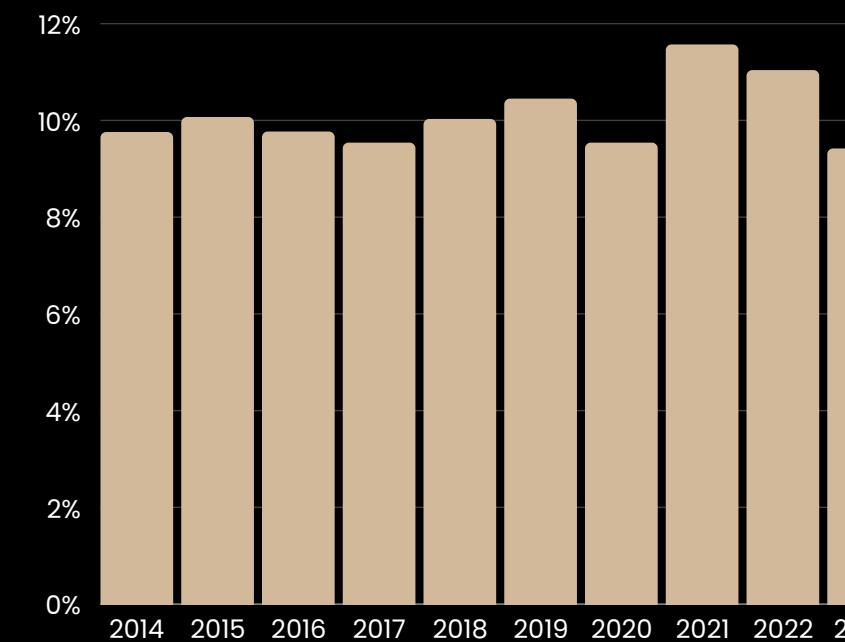
# STRATEGIA TIME – PENULTIMO

5 HIT RATE

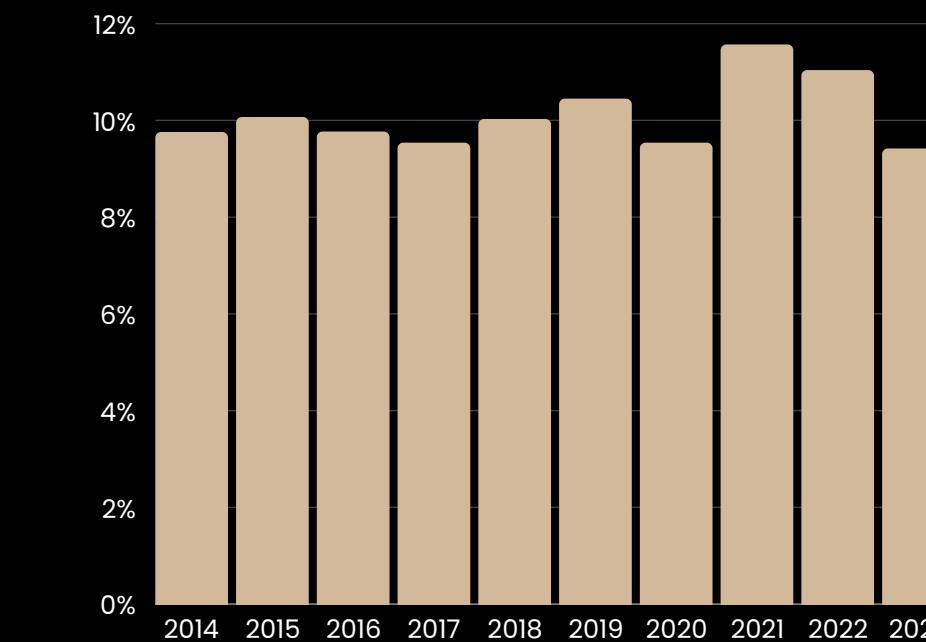
Llama3.1 8b



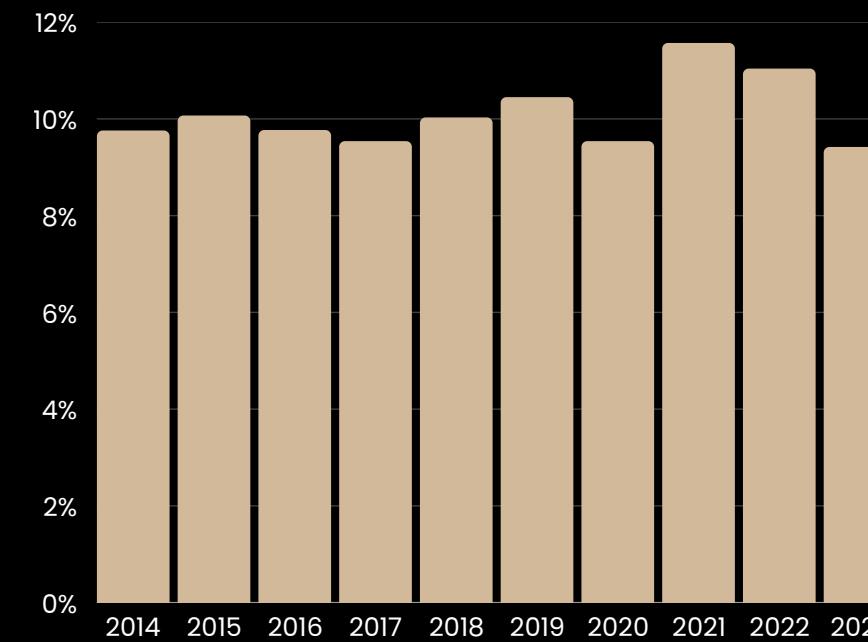
Qwen2.5 7b



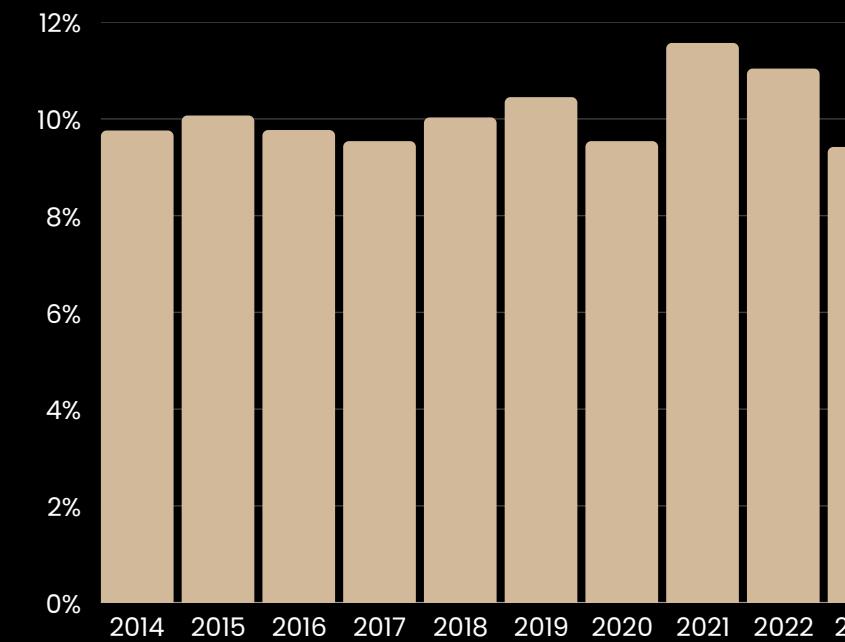
Qwen2.5 14b



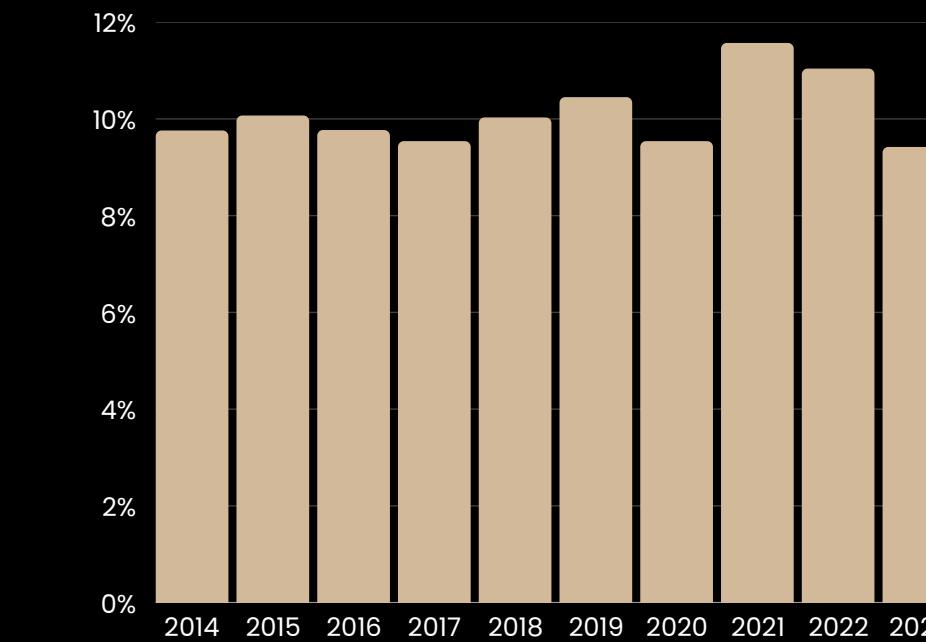
Mixtral 8x7b



Mistral 7b



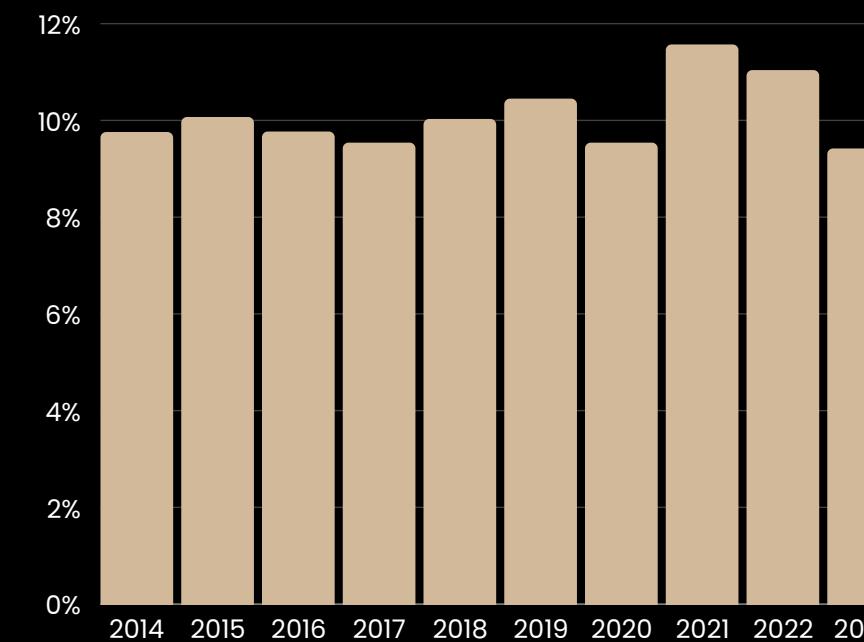
Deepseek Coder 33b



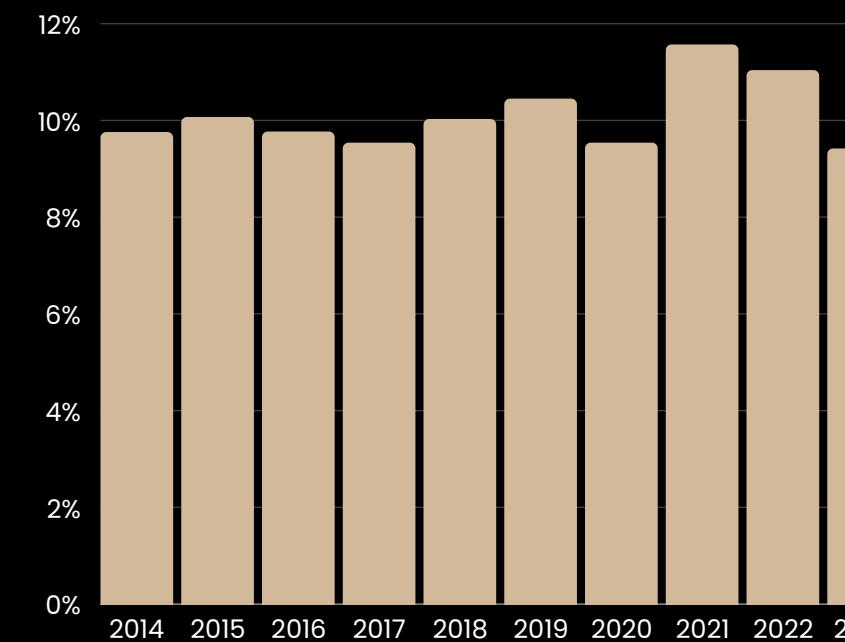
# STRATEGIA TIME - METÀ

TOP 1 ACCURACY

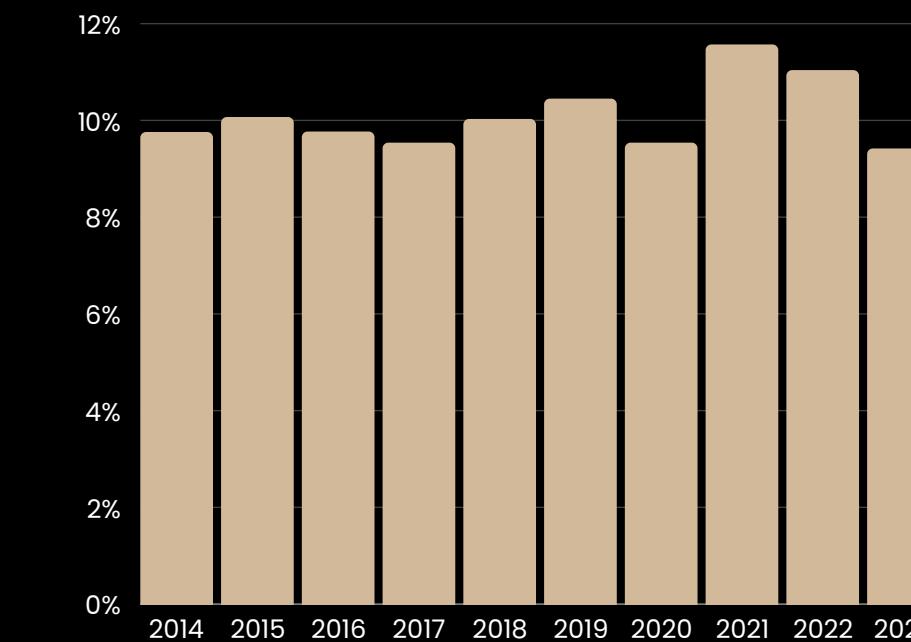
Llama3.1 8b



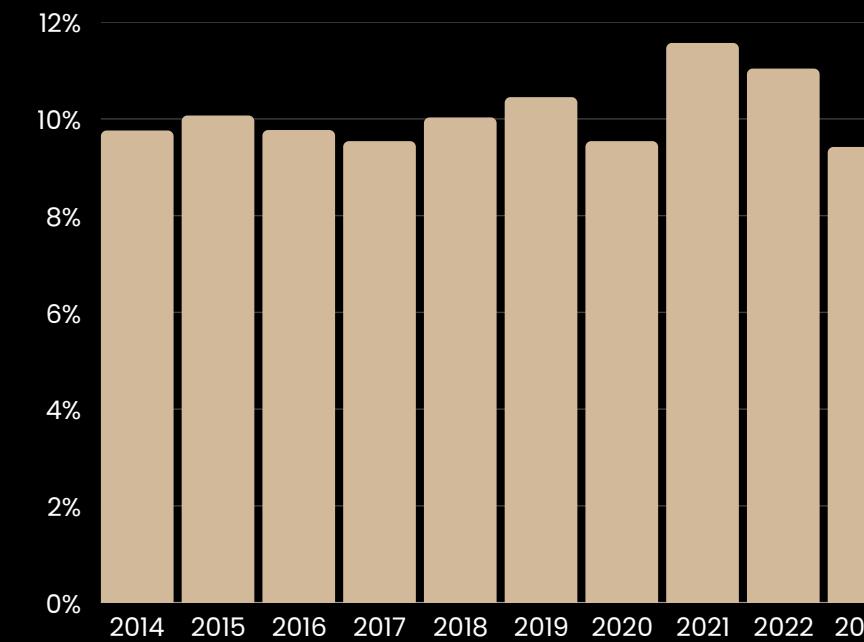
Qwen2.5 7b



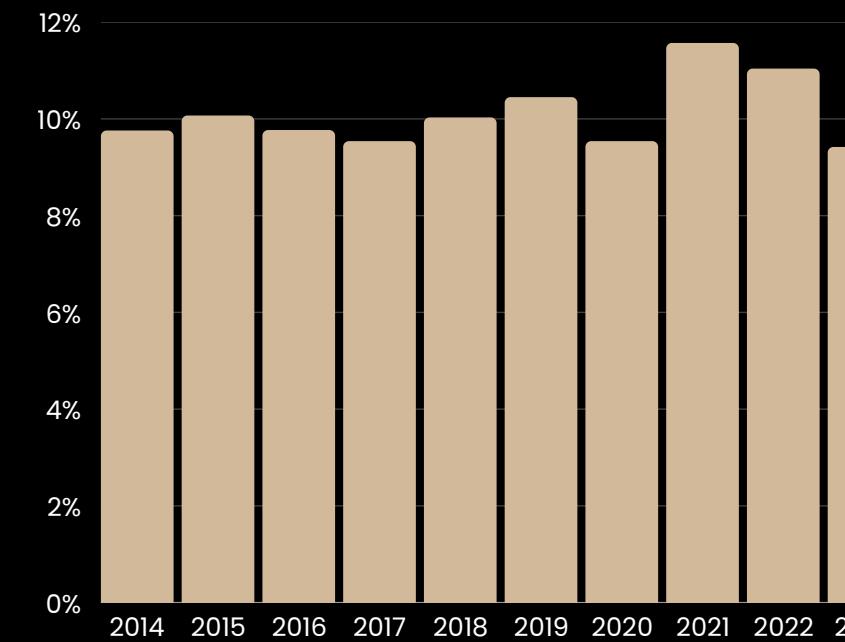
Qwen2.5 14b



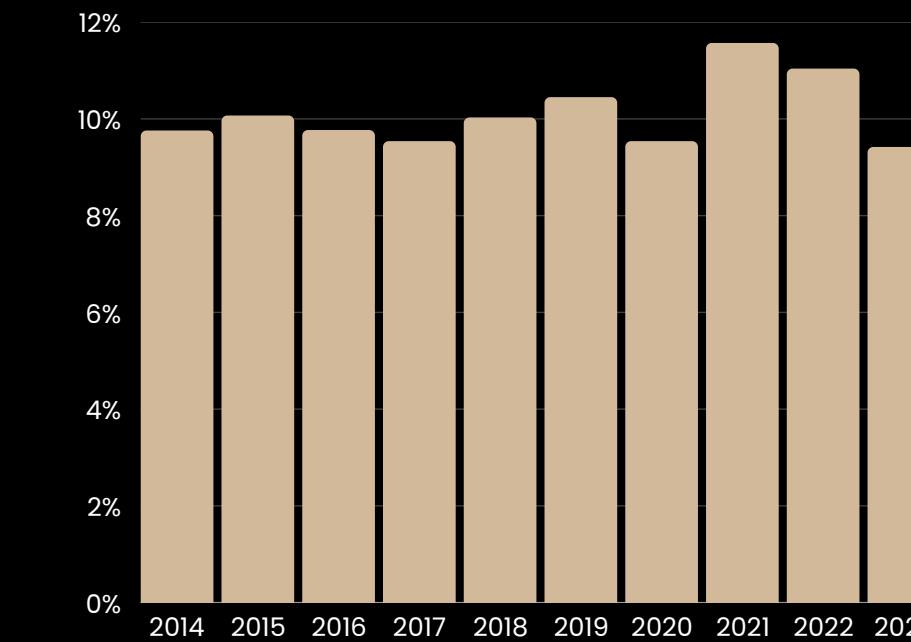
Mixtral 8x7b



Mistral 7b

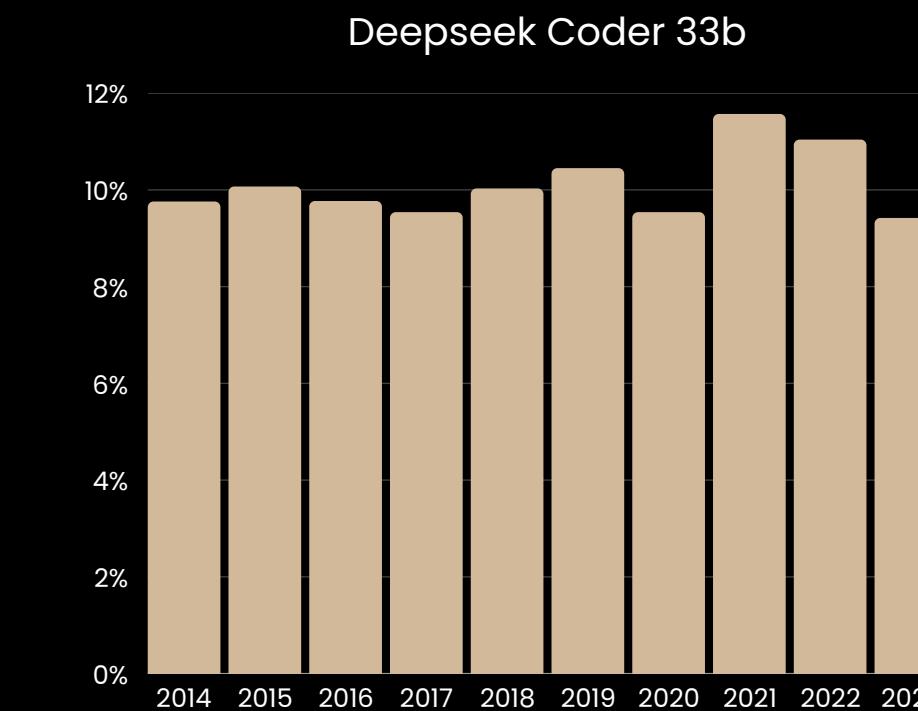
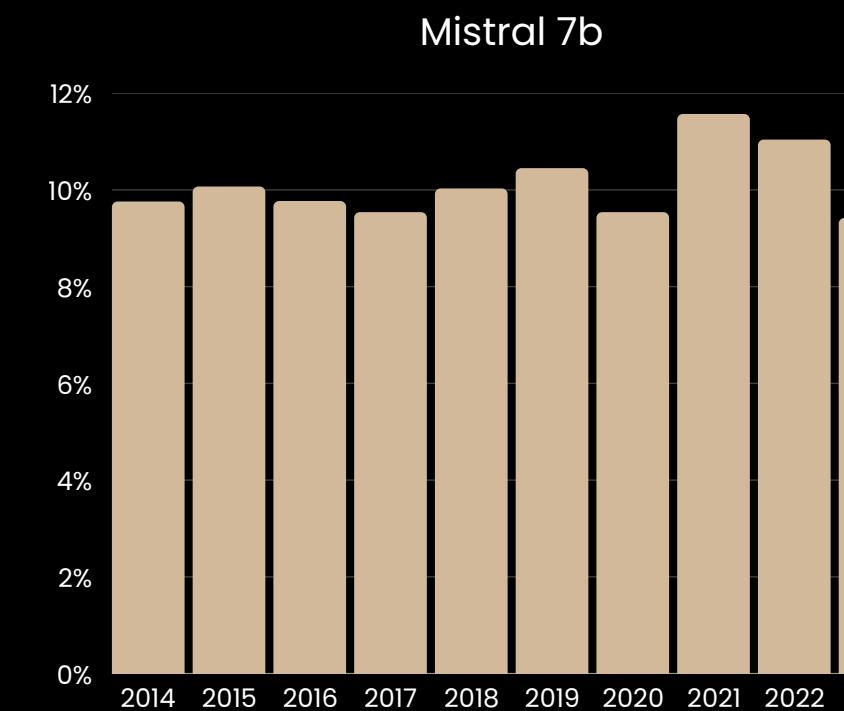
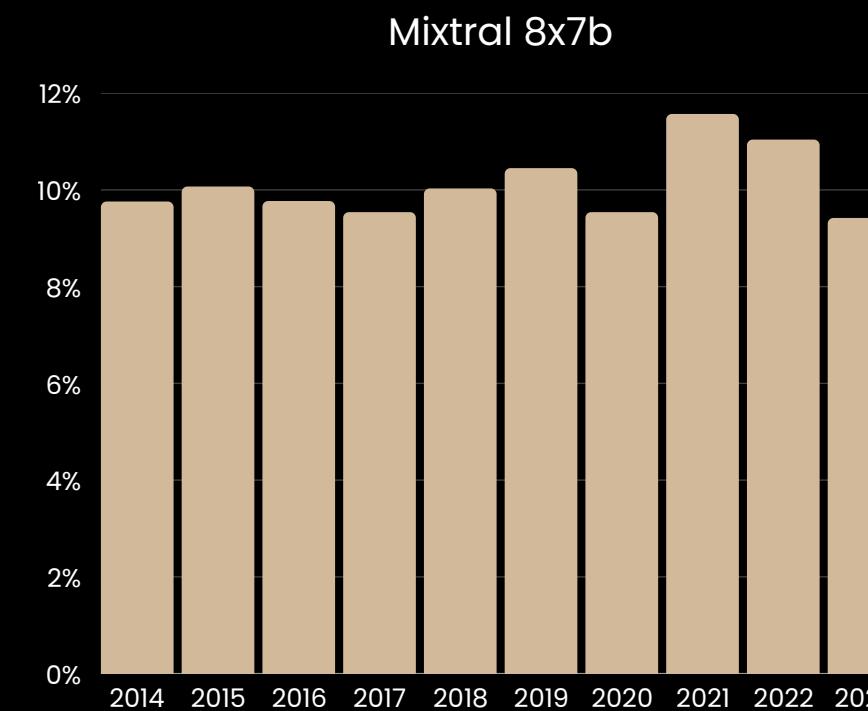
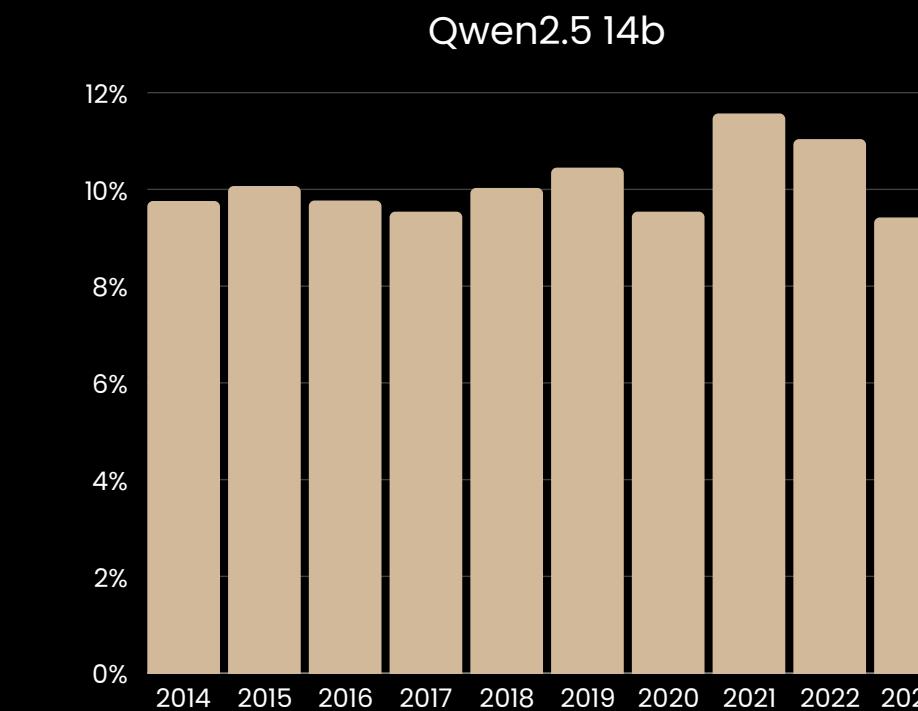
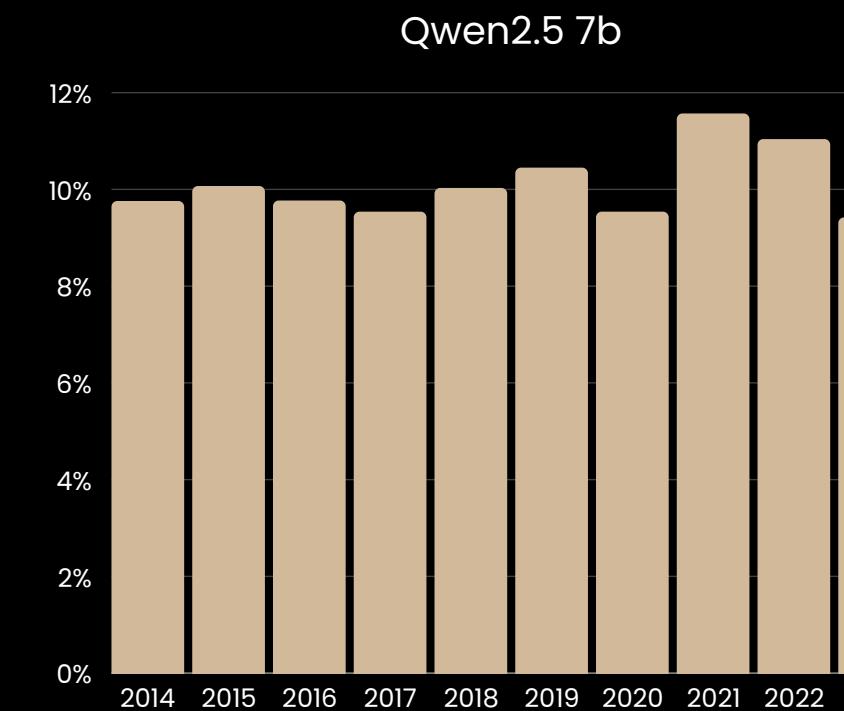
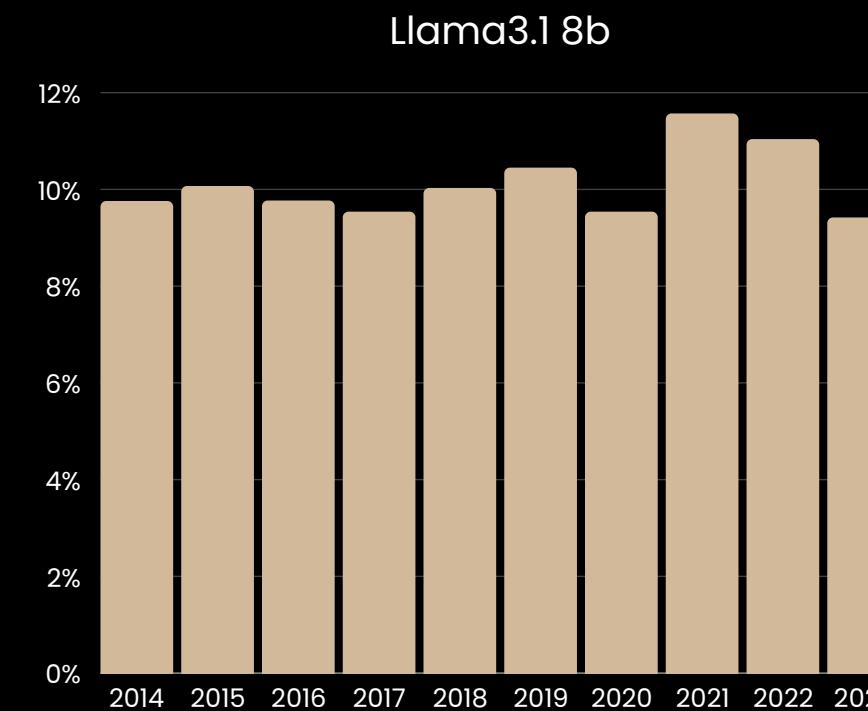


Deepseek Coder 33b



# STRATEGIA TIME - METÀ

5 HIT RATE

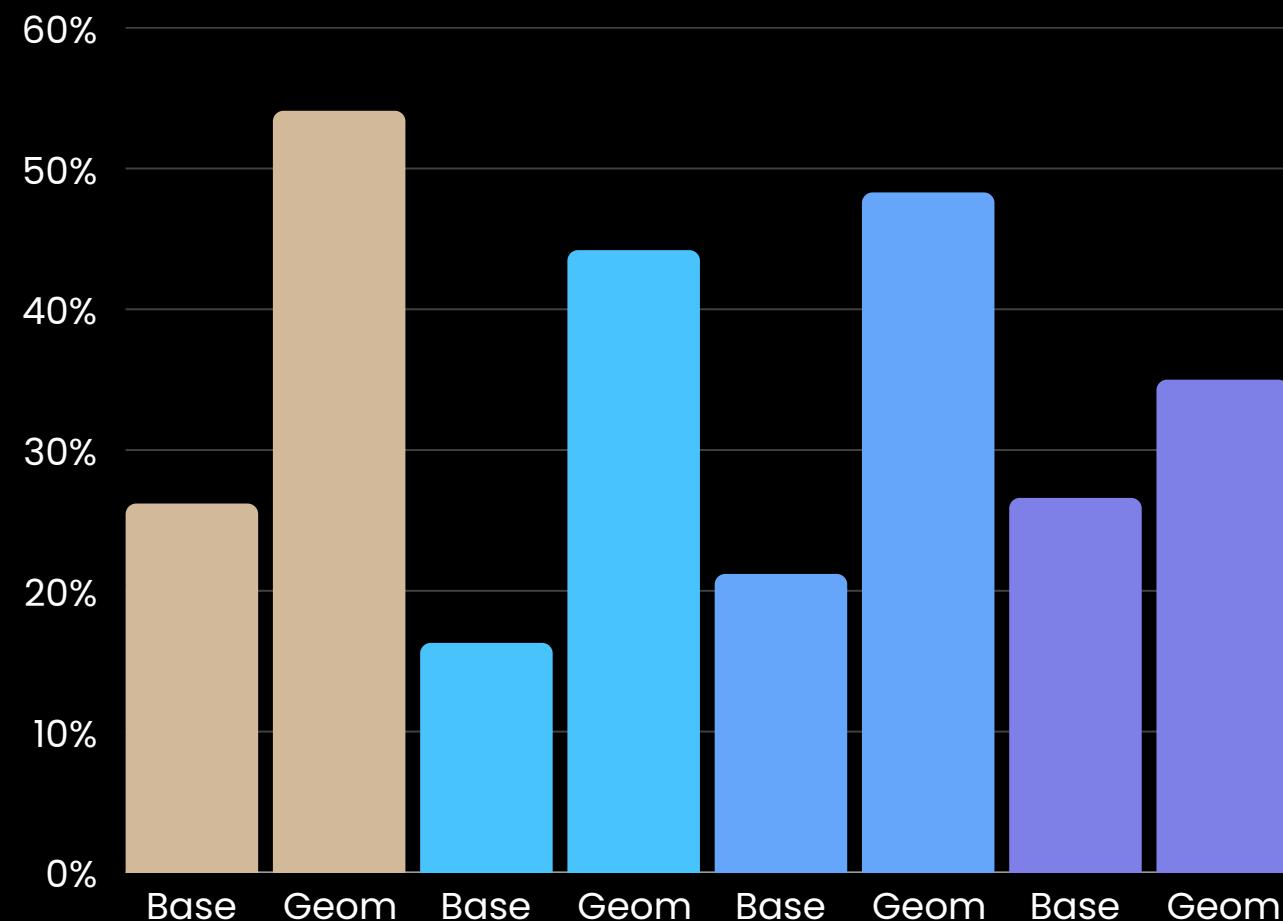


# CONCLUSIONE

---

08

# CONFRONTO INTER-MODELLO



## Llama3.1 8B

Our domestic sales have remained stable this year, reflecting our strategic efforts to maintain market presence and improve customer satisfaction.

## Mixtral 8x7B

International sales begin to increase in the second quarter, driven by expanding into new markets and strengthening partnerships with key distributors.

Qwen2.5 14B

Qwen2.5 7B

Deepseek Coder 33B

# IMPLEMENTAZIONI FUTURE

---

## SISTEMA REAL TIME NELLE CITTÀ

Per distribuire uniformemente i turisti nei punti di interesse

## GESTORE FLUSSI IMPIANTI SCIISTICI

Ottimizzazione distribuzione sciatori per evitare sovraffollamento

## MIGLIORAMENTO DEL CONTESTO

Ad Esempio:

- Meteo
- Calendario



# SUPPORTO PER QUESTO IL LAVORO

---



## CINECA

Accesso all'infrastruttura HPC Leonardo



## UNIVERSITÀ DI VERONA

Accesso Al DataSet VeronaCard



## PROF.SSA MIGLIORINI

Supervisione e guida durante lo sviluppo



## AMICI E COLLEGHI

Sostegno nei momenti più critici



## FAMIGLIA

Supporto durante l'intero percorso universitario

*"Se vuoi andare veloce, vai da solo. Se vuoi andare lontano, vai insieme"*

# GRAZIE PER L'ATTENZIONE

Telefono  
+39 3347985325

Email  
[simone.mattioli@studenti.univr.it](mailto:simone.mattioli@studenti.univr.it)

Dettagli Implementativi  
[Repository Git-Hub](#)