

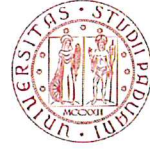


QUESITI COLLOQUIO - SELEZIONE 2025S48

Elenco n.1

1. Traduci questo testo (tratto da Ahlmann-Eltze, et al., 2025).

“In summary, we presented prediction tasks where current foundation models did not perform better than deliberately simplistic linear prediction models, despite significant computational expenses for fine-tuning the deep learning models. As our deliberately simple baselines are incapable of representing realistic biological complexity, yet were not outperformed by the foundation models, we conclude that the latter’s goal of providing a generalizable representation of cellular states and predicting the outcome of not-yet-performed experiments is still elusive.”
2. Descriva come l'integrazione di un sistema di controllo di versione distribuito (come Git) supporta lo sviluppo di pacchetti R.
3. Discuta il concetto di containerizzazione (ad esempio, tramite Docker) come approccio alla distribuzione del software in ambito scientifico.



Elenco n.2

1. Traduci questo testo (tratto da Ahlmann-Eltze, et al., 2025)
“Our results are in line with previously published benchmarks that assessed the performance of foundation models for other tasks and found negligible benefits compared to simpler approaches. Our results also concur with two previous studies showing that simple baselines outperform GEARS for predicting unseen single or double perturbations. Since the release of our paper as a preprint, several other benchmarks were released that also show that deep learning models struggle to outperform simple baselines. Two of these preprints suggested an even simpler baseline than our linear model, namely, to always predict the overall average, and we have included this idea here.”
2. Illustri i principi fondamentali e gli strumenti per la creazione e la manutenzione della documentazione dei pacchetti R.
3. Discuta la configurazione e l'utilizzo di ambienti virtuali di gestione del software e i loro vantaggi rispetto ad un'installazione globale a livello di sistema operativo.



Elenco n.3

1. Traduci questo testo (tratto da Ahlmann-Eltze, et al., 2025)
“One limitation of our benchmark is that we used only four datasets. Another limitation is that all datasets are from cancer cell lines, which, for example, Theodoris et al. excluded from their training data because of concerns about their high mutational burden. We also did not attempt to improve the original quality control, for example, by excluding perturbations that did not affect the expression of their own target gene and, thus, might not have worked as intended.”
2. Quali sono le principali metodologie di testing per assicurare la stabilità e la correttezza funzionale di un pacchetto R?
3. Analizzi le problematiche critiche relative alla gestione, all'installazione e all'aggiornamento di software scientifico e delle sue librerie in un contesto di sistema multiutente condiviso.