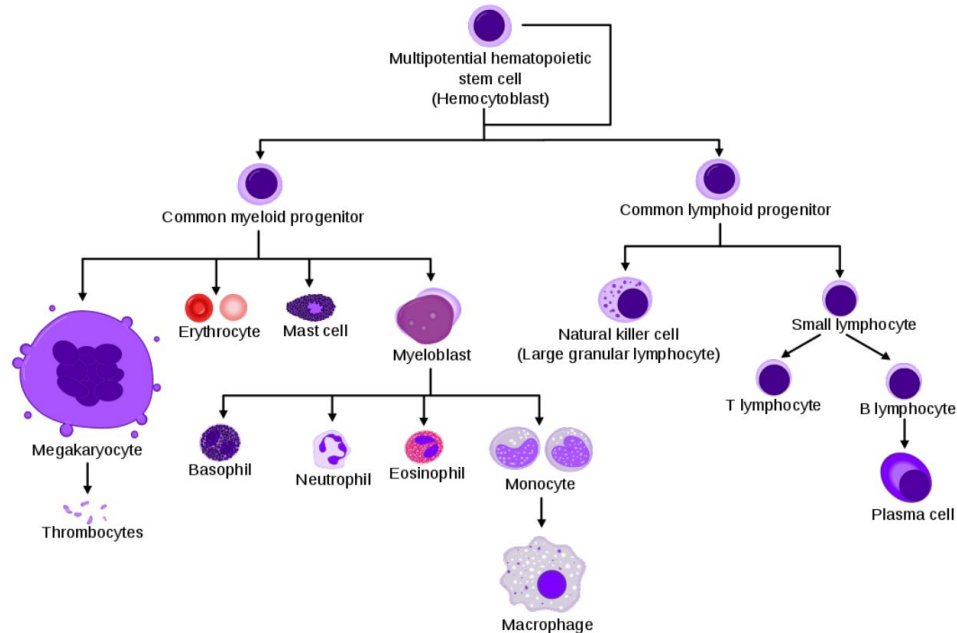


Hierarchical, multi-label prediction for automated cell type identification

Iwijn Voeten
Klaas Goethals

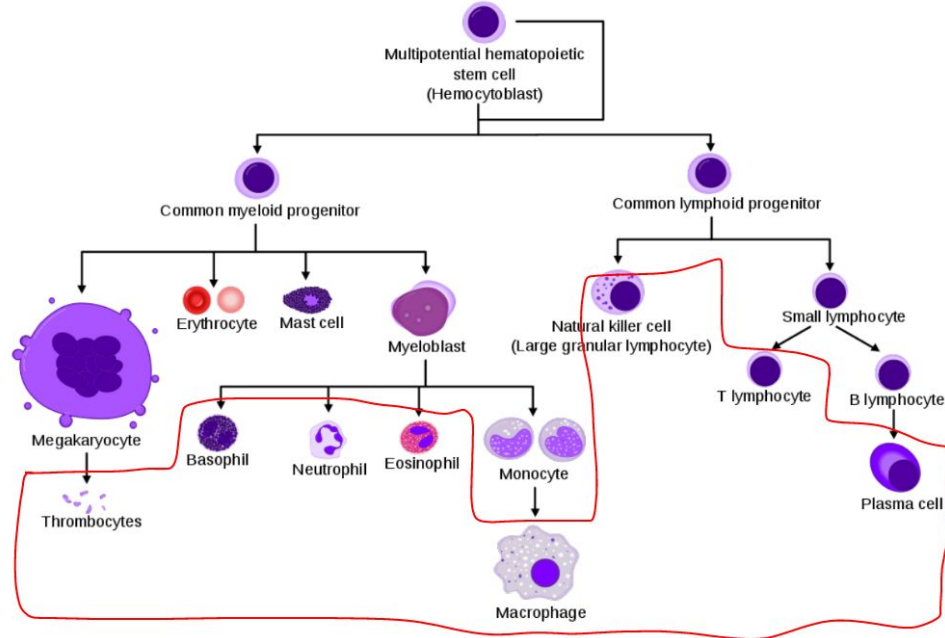
Problem description

- Cell types have hierarchical structure
- Multiple approaches to classifying cells



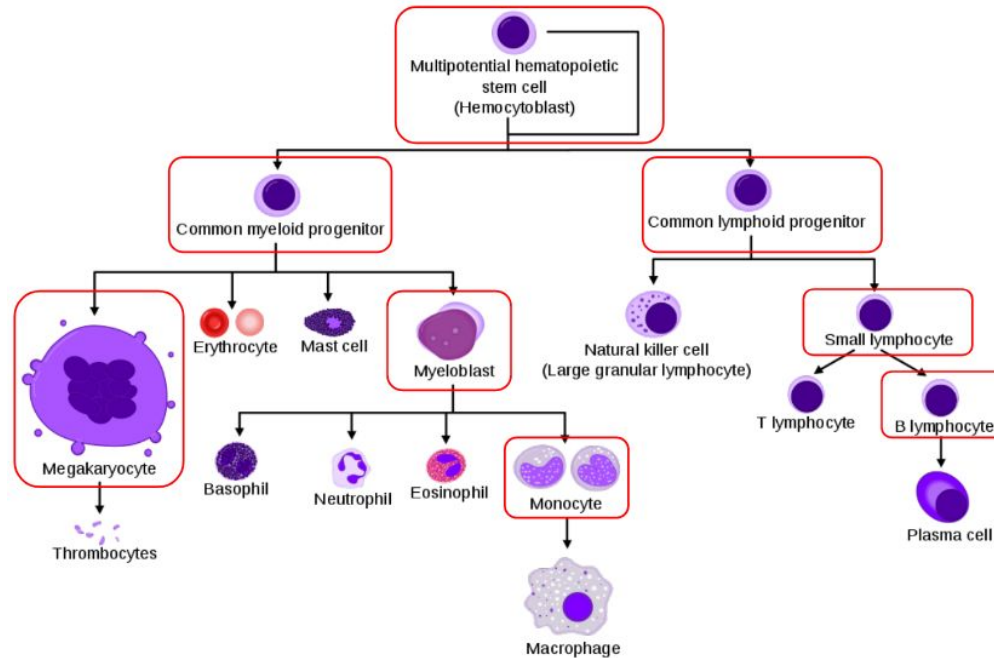
Different approaches

- Flat classifier



Different approaches

- Hierarchically-Structured Local Classifiers



Different approaches

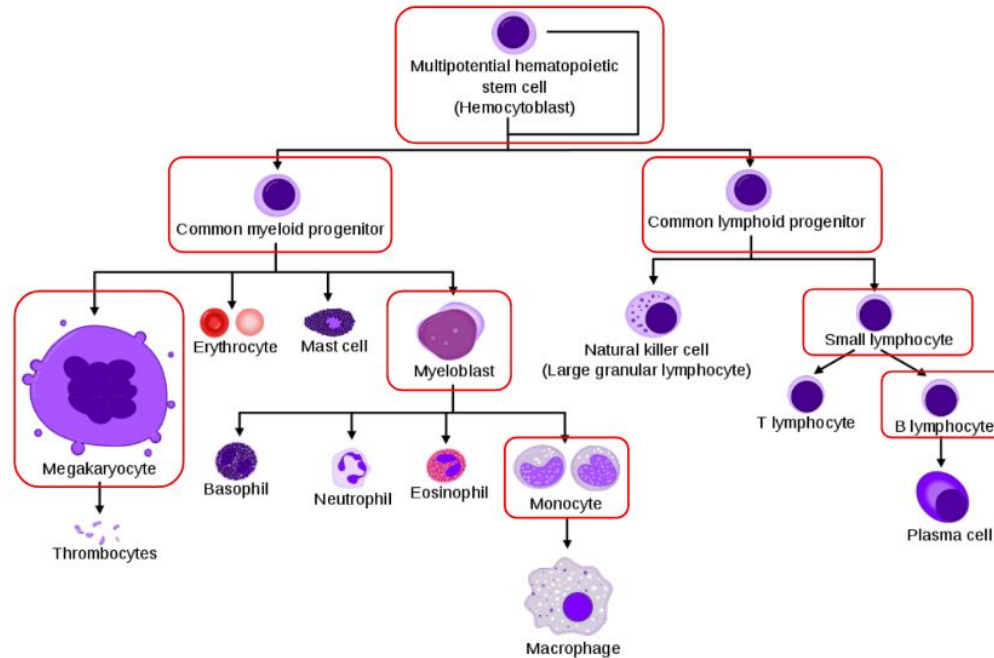
Hierarchically-Structured Local Classifiers

Multiple options:

- Local Classifier per Node (LCN)
- Local Classifier per Parent Node (LCPN)
- Local Classifier per Level (LCL)

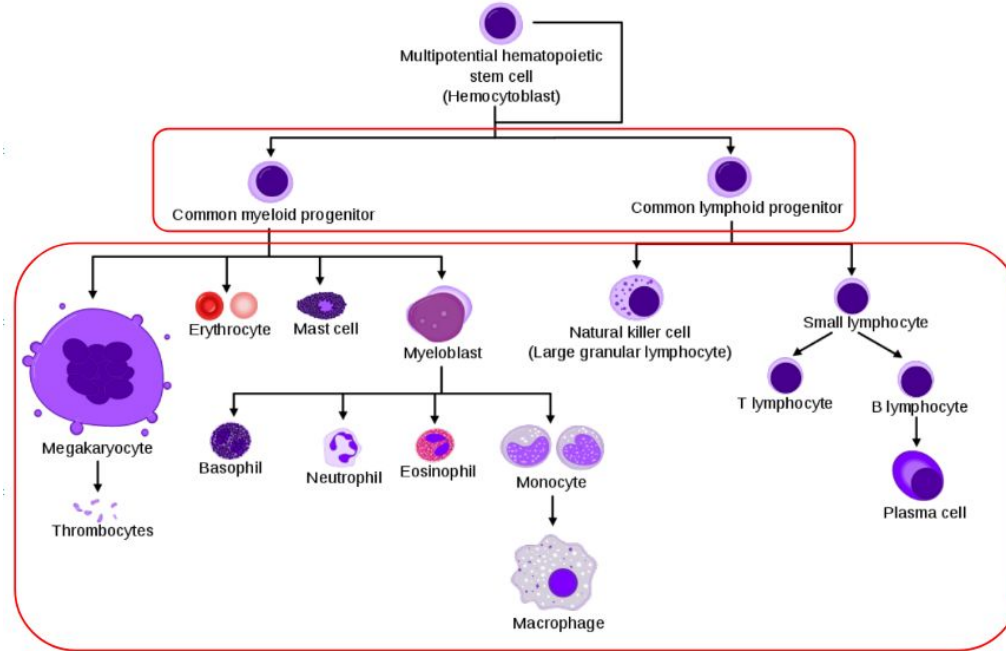
Different approaches: hierarchically-structured

- Local Classifier per Parent Node (LCPN)



Different approaches: hierarchically-structured

- Local Classifier per Level (LCL)



What we did

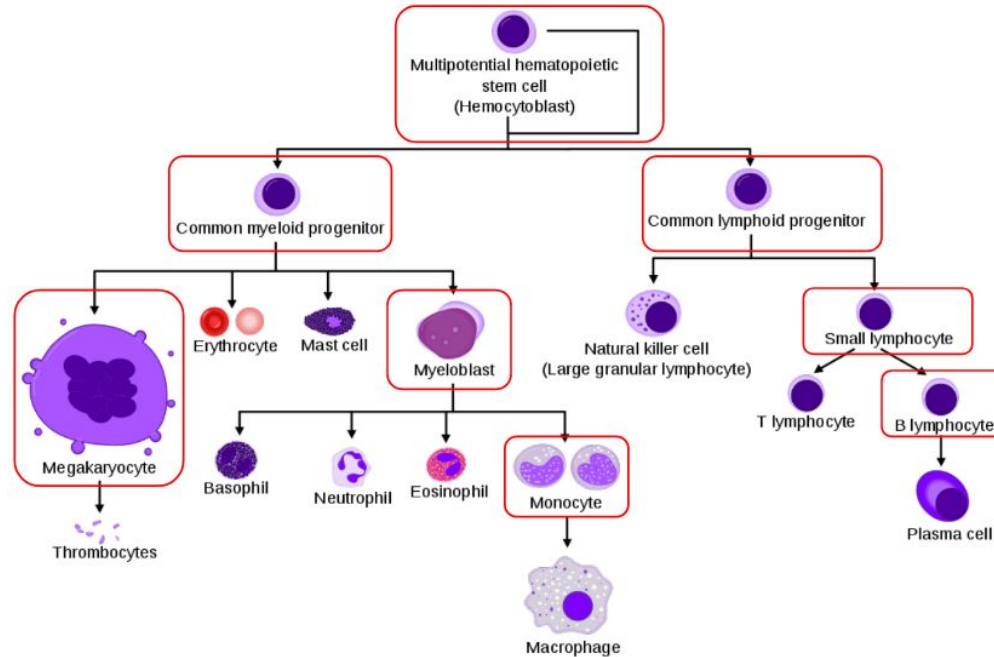
- Data filtering
- Flat classifier
- Hierarchical: Local Classifier per Parent Node (LCPN)
 - Limited depth
 - Depth all the way
- Hierarchical: Local Classifier per Level (LCL) modified

Data

- “Allen Mouse Brain” dataset
- “Covid BAL” dataset
- Remove columns consisting of all or many zeros

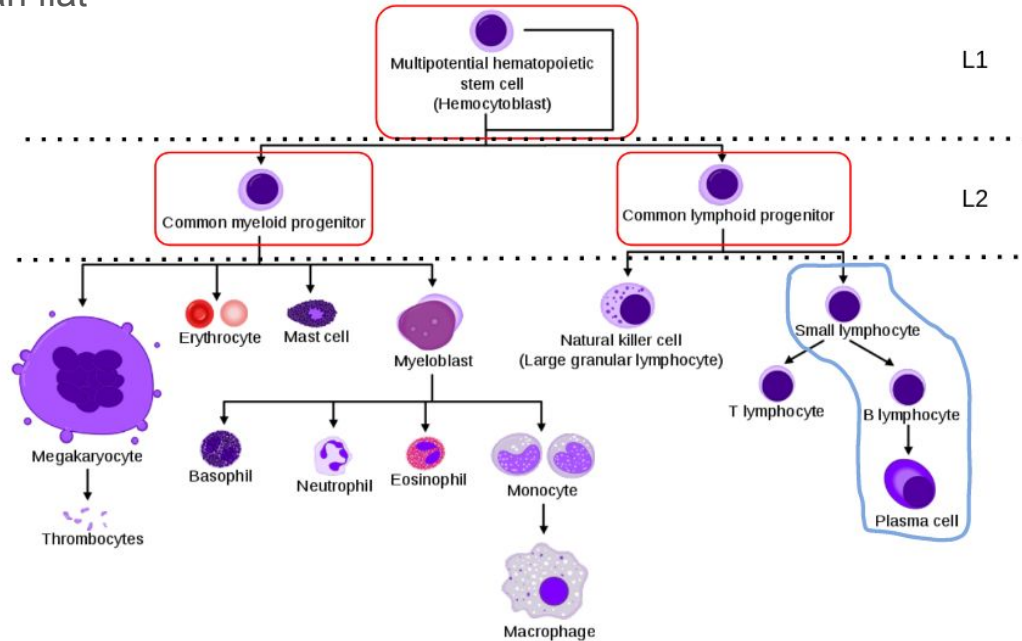
Flat vs Hierarchical

- Hierarchical performed worse



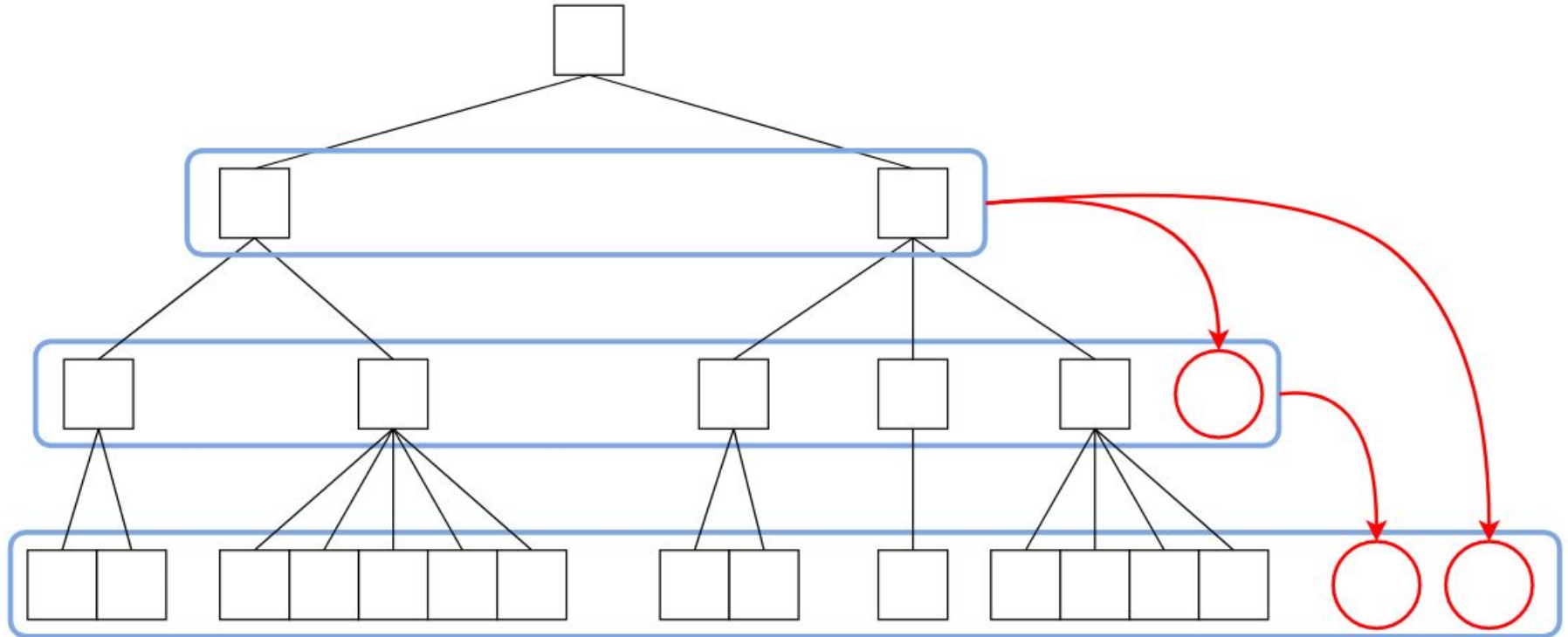
Limited Hierarchical

- Depth limiting
 - Better than full hierarchical
 - Faster than flat



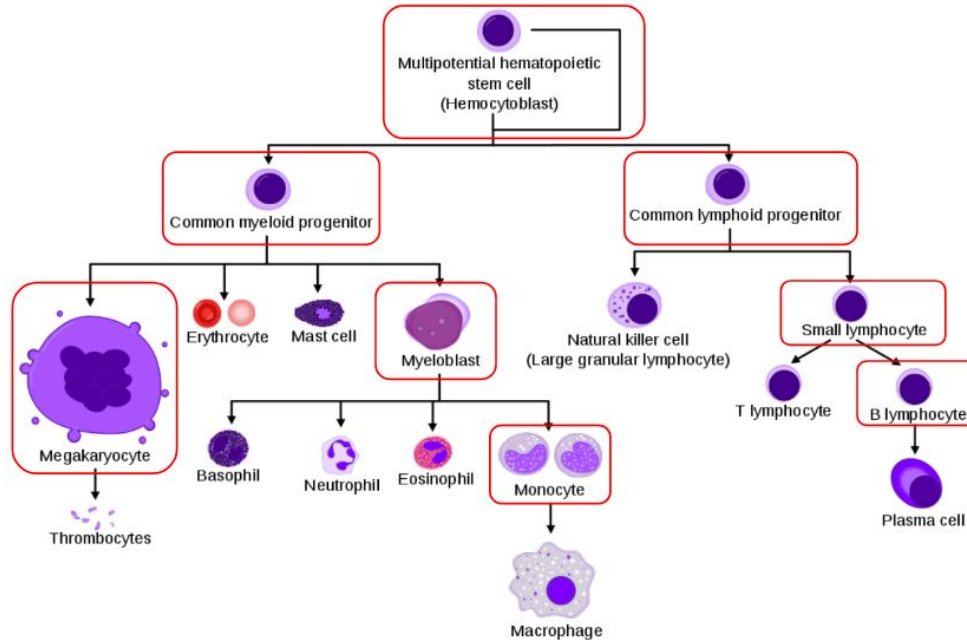
Local Classifier per Layer modified

- Correct mistakes made earlier



Parallelization hierarchic models

- Train models on multiple cores
- Unsuccessful



Results

- Time decrease of 18.8% with limited hierarchical model
- Time decrease of 21.6% with full hierarchical model

- Flat classifier: 90.2% accuracy
- Limited hierarchical: 90.2% accuracy
- Full hierarchical: 47.2% accuracy

Future work

- Hierarchical classifier: Local Classifier per Node
- Hierarchical limited on more depths