

Mode 1: Rayyan screening

Updated set (N=1,126)



Load in Rayyan



Manual screening



Relevant (+)



Irrelevant (-)

Mode 2: AIM Review screening

Updated set (N=1,126)



Load in AIM Review



Manual screening



Relevant (+)



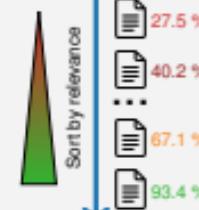
Irrelevant (-)

Mode 3: AIM Review screening with active learning

Updated set (N=1,126)



Predict relevance scores



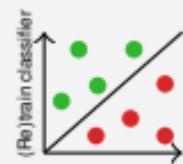
Manual screening



Relevant (+)



Irrelevant (-)



Original set (N=3,118)

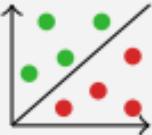
Mode 4: AIM Review automatic screening

Relevant (+)

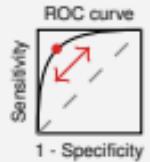


Irrelevant (-)

Nested cross-validated classifier development



Probability threshold selection



Updated set (N=1,126)



Predict unlabelled