Ranking hypothesis:

1. Ranking should value search input of the user the most - this input contains both search keyword and location
2. We should significantly derank ineligible and throttled jobs
3. Title/ sub department matching of the jobs should be prioritised - but not at the cost of deranking search input based results
4. Preference mismatch jobs should be deranked for the user - but not at the cost of deranking search input based results
5. Match score of the job should be prioritised - but not at the cost of deranking search input based relevance
6. In case of area based searches - area input of the user should be significantly prioritised (since the user has shown intent to apply for area specific jobs)

Case 1: User enters a keyword based search query + city

i.e. custom search is not done and area input is not provided

e.g. search query is *"Primary teacher"*

Term match	High match score
e.g. "Primary teacher"	Sub-dept/ Title match
	Others
	Preference mismatch
Phrase match	High match score
e.g. "Pre primary teacher"	Sub-dept/ Title match

	Others
	Preference mismatch
Title affinity	High match score
e.g. "School Teacher"	Sub-dept/ Title match
	Others
	Preference mismatch
Partial match	High match score
e.g. "teacher"	Sub-dept/ Title match
	Others
	Preference mismatch
Ineligible jobs	
Throttled jobs	

- ES score in these cases ranges from 40 to 2K+
- The buckets can be characterised by ES score
  - 1K+ ES score
    - Typically for term matches and phrase matches
  - 200-1K ES score
    - Typically for similar title matches
  - 40-200 ES score

        - Typically for partial matches

Case 2: User performs a CustomSearch

i.e. WFH, nightshift etc. there is no concept of keyword based matching here hence, ES score is not passed

High match score
Sub-dept/ Title match
Others
Preference mismatch
Ineligible jobs
Throttled jobs

Case 3: User performs a keyword based or custom search + AREA

- Split into:
    - Area match
    - Non area match

Area match bucket	High match score + ES context
	Sub-dept/ Title match + ES context
	Others + ES context

	Preference mismatch + ES context
Non area match bucket	High match score + ES context
	Sub-dept/ Title match + ES context
	Others + ES context
	Preference mismatch + ES context
Ineligible jobs	
Throttled jobs	

Detailed ranking logic:

Examples:

https://docs.google.com/spreadsheets/d/1PoSkBSJysf8_nX7DApSW-mA6zrdj6G7PXipjNdSJBCg/edit?usp=sharingConnect your Google account

Step 1: es_normalised_score

ES score normalisation

- If es_score > T1
  - Then 1
  - Else es_score/TI
- Here, T1 is the ES score threshold

- We will not be doing more detailed ES score normalisation since it is out of scope for this ranking exercise which is to optimise for recruiter and candidate experience - AND NOT stronger search query context for the user

Step 2: adjusted_es_score

Scoring inclusion for: Prioritisation for match score and Sub-dept/ title match &

Deprioritisation basis preference mismatch score

- if match_score > T2 i.e. match_score cutoff
    - Then (es_normalised_score*W1+match_score*W2+preference_mismatch_score*W4)/(W1+W2+W4)
    - Else
        - If sub_dep_title_match = 1
            - Then (es_normalised_score*W1+sub_dep_title_match*W3+preference_mismatch_score*W4)/(W1+W3+W4)
            - Else (es_normalised_score*W1+preference_mismatch_score*W4)/(W1+W4)
    - Here - es_normalised_score comes from step 1
    - match_score is pulled from module 3
    - sub_dep_title_match
        - Would be boolean
        - Would have to be extracted from dealbreaker eligibility score component
            - Note that in it, mismatch would be 1. We would have to make changes accordingly
        - is pulled from module 3
    - preference_mismatch_score
        - Dealbreaker preference score component
        - is pulled from module 3

Step 3: net_score

Significant Deprioritisation for eligibility mismatch, throttled job

adjusted_es_score+W5*revised_eligibility_mismatch_score+W6*throttling

- adjusted_es_score from step 2
- revised_eligibility_mismatch_score
    - Should be 0-1
    - Would have to be extracted from dealbreaker eligibility score component MINUS sub_dep_title_match component
    - W5 and W6 would be negative values

Weights and cutoff values for initialisation:

T1	ES Score cutoff	2000
W1	ES weight	2
W2	Match score prioritisation	3
T2	Match score cutoff	0.6
W3	Sub-department/ title afifnity match prioritisation	1
W4	Preference mismatch score depriorisation	0.1
W5	Eligibility mismatch score depriorisation	5
W6	Throttled job depriorisation	10

Handling edge cases:

1. **User performs a CustomSearch (8% cases)**
    1. Follow the same scoring logic
    2. Since no es_score would exist here, start from step 2 and remove the es_normalised_score component
2. **User performs area based search (12% cases)**
    1. Introduce a bucket called “Area match”
        1. Follow the same scoring logic, to net_score add **+ 1 for area matches**
    2. For non “area matches”,
        1. Follow the same scoring logic

Rollout conditions:

- This would be rolled out as a child cohort of 0 eligibility logics on search

What problem are we solving?

- We are currently using the older taxonomy which is outdated and not being maintained
- Thus the logics leveraging the same are giving fewer/ erroneous results
- Also the affinity logics used across entities needs to be revisited to give higher weightage to relevant entities only e.g., reduce skills based logics

What are the hypothesis to solve the problem? Approach?

- Migrate all taxonomy logics to taxonomy 2.0 vector based logic which includes
    - Main buckets pulling logic for suggested entity
    - Other bucket pulling logic for suggested entity
    - Scoring service changes
- The scoring service changes include match score, preference dealbreaker and eligibility dealbreaker which should be brought at par with feed handling e.g., specialisation & degree score leveraging the new taxonomy. Not required post “Scoring profile integration change”
- For the pull logic on Main & Other bucket please go through the details below

- Note here the identified title entity including search entity and affinity/ similar entities are searched across Job title, dub-dep, dep, org, skills, description, the priority order of which is defined basis score below which needs to be updated
- Further the identified entities are search considering exact keyword match, phrase match, individual word match (except stop words) etc.

Entity	Current score	New proposed score
Job title	10	10
Job sub-dep	3	2
Job department	3	1
Job organisation	5	3
Job skills	3	0.1
Job description	1	0.5

Search Entity type in order of usage (Ref)	Bucket	Mapping utilised	Condition	Ranking consideration	Comments

Search Entity type in order of usage (Ref)	Bucket	Mapping utilised	Condition	Ranking consideration	Comments

Title	Main	Affinity Titles	>=X1	Vector score	
	Other	Affinity Titles	<X1, >=X2	Vector score	
		Affinity Sub-dep	>=X3	Vector score	Very high sub-dep match to be only included here
Raw Text/ Raw Suggestions ?	Main	Search the entered keyword only No mapping utilised	-	-	
	Other	No mapping utilised Job Feed jobs	-	-	Name the section as jobs based on your profile
Custom Search	Main	Search results to be filtered jobs list basis custom filter	-	-	

	Other	No mapping utilised Job Feed jobs	-	-	Name the section as jobs based on your profile
Organisation	Main	Exact organisation match (Remove stop words)	?	?	?
	Other	No mapping available in system If industry for company is available/ accessible show all jobs from same industry else show Job Feed jobs			If Job Feed jobs are being shown then name the section as jobs based on your profile
Sub-department	Main	Affinity Titles	>=X4	Vector score	
		Affinity Sub-dep	>=X5	Vector score	

	Other	Affinity Titles	<X4, >=X6	Vector score
		Affinity Sub-dep	<X5, >=X7	Vector score
Department	Main	~~Affinity Titles~~	~~>=X8~~	~~Vector score~~
		~~Affinity Sub-dep~~	~~>=X9~~	~~Vector score~~
		Affinity Department	>=X10	Vector score
	Other	~~Affinity Titles~~	~~<X8, >=X11~~	~~Vector score~~
		~~Affinity Sub-dep~~	~~<X9, >=X12~~	~~Vector score~~
		Affinity Department	<X10, >=13	Vector score
Skill	Main	Affinity Titles	>=X14	Vector score
		Affinity skills	>=X15	Vector score

Other	Affinity Titles	<X14, >=X16	Vector score
	Affinity skills	<X15, >=X17	Vector score

1. X1, X2, X3….Xi are configurable thresholds