Entity Resolution and Link Prediction

Link prediction and entity resolution are two ways to identify missing information in networks. Link prediction helps identify edges that are likely to appear in the future, if they do not exist already. Entity resolution uses attributes and network structure data to link nodes that represent the same individual.

Entity resolution

Entity resolution is the process that resolves entities and detects relationships. The pipelines perform entity resolution as they process incoming identity records in three phases: recognize, resolve, and relate.

Recognize: During entity resolution, pipelines must recognize the data by validating, optimizing, and enhancing the incoming identity data. During this recognize phase of the pipeline process, the pipelines cleanse and standardize the data values, as well as perform data quality checks on the data to protect the integrity of the entity database.

Resolve: During entity resolution, the pipelines resolve identities into entities. After the data values in the identity records have been cleansed, standardized or enhanced, the pipeline uses sophisticated search algorithms to compare the data values in the incoming identity record against existing entities in the entity database to determine if they are the same entity.

Relate: During entity resolution, pipelines also complete the relationship detection process, which detects relationships between identities and entities and generates alerts for relationships of interest.

Scoring: During entity resolution, the system computes how closely the attributes for an incoming identity match the attributes of an existing entity. The results of this computational analysis are scores that the system uses to resolve identities into entities and detect relationships between entities.