Appendix B
Baseline Approaches

We compare `iFit` with the following eight baseline approaches in the literature [14]:

- `h-index` [9]: It calculates the h-index of a class in the software network as its importance. Generally, if a class has N_p neighbors that are ranked in descending order according to their received degrees, then the class has h-index h if each top-h neighbor receives a degree $\geq h$, and the other $(N_p - h)$ neighbors receive a degree $< h$ each. The top-h neighbors form the Hirsch core.
- `a-index` [9]: It calculates the a-index of a class in the software network as its importance. Formally, $a = \frac{1}{h}\sum_{j=1}^{h} d_j$, where a is the a-index value, h is the h-index value, d_j is the degree of class j in the Hirsch core (cf. `h-index`), and the h classes in the Hirsch core are ranked in descending order according to their received degrees.
- `k-core` [10]: It applies a k-core decomposition to calculate the *coreness* of a class in the software network as its importance. The k-core of a network is defined as the largest subgraph, where each node has a degree $\geq k$. Then a node has coreness k if it belongs to the k-core but not to the $(k+1)$-core.
- `PageRank` [19]: It applies a PageRank algorithm to calculate the PageRank value of a class in the software network as its importance. The PageRank value of class u (i.e., $PR(u)$) is computed by $PR(u) = \frac{1-d}{m} + d \times \sum_{v \in IN(u)} \frac{PR(v) \times w(v,u)}{wOutDeg(v)}$, where $IN(u)$ is the in-neighbor of class u, $w(v,u)$ is the weight on the link from v to u, $wOutDeg(v)$ is the weighted out-degree of class v, and m is the number of classes in the software network. d is the damping factor which is fixed to 0.85.
- `PageRank_BR` [19]: It improves the `PageRank` by considering the back recommendation. Specifically, if there is a link from classes `A` to `B`, the authors treated it as a forward recommendation of `A` on `B`. For each forward recommendation, they added a back recommendation (i.e., a link from `B` to `A`) with the weight on the link being a fraction F of the weight on the corresponding forward recommendation. In this work, we used the best value of F reported in [19], i.e., F=0.5.
- `ICOOK` [11]: It calculates the *generalized coreness* of a class in the software network as its importance. The *generalized coreness* is computed on weighted directed networks and based on generalized degrees.
- `ElementRank` [12]: It takes a similar way as `PageRank` to compute the PageRank value of classes. The only difference is that $\frac{1-d}{m}$ is replaced by $\frac{(1-d) \times InLinks(u)}{\sum_{k=1}^{N} InLinks(k)}$, where $InLinks(u)$ is the sum of weights over the incoming links of node u.
- `Pride` [14]: It takes a similar way as `PageRank` to compute the PageRank value of classes. The only difference is that $\frac{1-d}{m}$ is replaced by $\frac{(1-d) \times wDeg(u) \times Deg(u)}{wDegSum \times DegSum}$, where $wDeg(v)$ is the weighted degree of class v, $Deg(v)$ is the degree of class v, $wDegSum$ is the sum of the weighted degree of all classes, and $DegSum$ is the sum of the degree of all classes.