

Terdapat dua faktor yang mempengaruhi minat seseorang untuk melanjutkan studi ke perguruan tinggi dapat berasal dari faktor internal (fisiologis, dan psikologis) ataupun *external* (sosial, dan non sosial).

Berdasarkan faktor-faktor ini disertai dengan data yang tepat, dapat dibangun sebuah sistem yang mampu memprediksi keputusan dari siswa. Dengan penerapan data mining, faktor-faktor yang mempengaruhi ketidakmampuan siswa melanjutkan ke perguruan tinggi dapat diperbaiki dan dicari solusinya. Dengan harapan meningkatkan jumlah sarjana dalam

Berdasarkan faktor-faktor ini disertai dengan data yang tepat, dapat dibangun sebuah sistem yang mampu memprediksi keputusan dari siswa. Dengan penerapan data mining, faktor-faktor yang mempengaruhi ketidakmampuan siswa melanjutkan ke perguruan tinggi dapat diperbaiki dan dicari solusinya. Dengan harapan meningkatkan jumlah sarjana dalam upaya mencerdaskan kehidupan bangsa.

Gambar 1.1 Ikan Goreng

Dataset dirancang secara artifisial menggunakan *package* sklearn pada python menggunakan fungsi `make_classification`. Data yang dihasilkan umumnya bersifat numerik dan memerlukan proses klasterisasi untuk mendapatkan data *nominal*. Pada data ini didefinisikan 10 kolom fitur dan 1 kolom label. Selain itu juga dibuat 1000 baris data dengan sebaran 50% untuk label “False” dan 50% untuk label “True”. Maksud dari “True” adalah siswa melanjutkan ke perguruan tinggi sedangkan “False” berarti tidak melanjutkan ke perguruan tinggi. Berikut adalah fitur-fitur yang ada pada dataset: