

ANALISIS PREDIKSI PROFIT PROYEK PEKERJAAN PLAFON GYPSUM MENGGUNAKAN KLASIFIKASI DECISION TREE

Kautsar Baihaqi Supriyadi, Feri Sulianta

Departemen Informatika, Universitas Widyatama, Indonesia

Email: kautsar.baihaqi@widyatama.ac.id, feri.sulianta@widyatama.ac.id

Abstrak

Perusahaan yang bergerak dalam bidang konstruksi yang salah satu layanannya yakni jasa pemasangan plafon gypsum, membutuhkan strategi kelola perencanaan anggaran atau Rincian Anggaran Biaya (RAB) untuk mendukung kelancaran pengerjaan proyek. Pada proses berjalannya proyek pekerjaan plafon, perencanaan anggaran digunakan sebagai sumber informasi utama dalam menghitung keuntungan perusahaan. Rincian Anggaran Biaya suatu proyek pekerjaan plafon gypsum meliputi biaya tenaga kerja, tenaga ahli, material, operasional proyek, biaya peralatan kerja dan biaya mobilisasi. Perencanaan yang tidak tepat dapat mempengaruhi cash flow sebuah perusahaan sehingga berkaitan erat terhadap keberlangsungan pelaksanaan proyek pekerjaan plafon. Diperlukan suatu teknik prediksi untuk menentukan keuntungan perusahaan berdasarkan penggunaan RAB suatu proyek agar perusahaan tersebut dapat menjaga cash flow yang mereka miliki sehingga tidak mempengaruhi pelaksanaan proyek. Tujuan untuk mencari faktor mana yang lebih menentukan untuk menghasilkan keuntungan dalam suatu proyek pekerjaan plafon. Penelitian ini menyajikan implementasi data mining dengan metode klasifikasi khususnya decision tree dapat menyelesaikan permasalahan tersebut dan menghasilkan sebuah knowledge discovery dalam membantu perusahaan mendapatkan keuntungan di masa depan. Kesimpulan, teknik klasifikasi dapat membantu perusahaan untuk lebih memperhatikan atribut tertentu dalam pelaksanaan sebuah proyek pekerjaan plafon dan membantu perusahaan atau pemangku kebijakan dalam menentukan sebuah keputusan di masa depan.

Kata kunci: Data Mining; Klasifikasi; Decision Tree; Prediksi; Rincian Anggaran Biaya; Profit.

Abstract

A company engaged in the construction sector, one of whose services is the installation of gypsum ceiling services, requires a strategy for managing budget planning or Detailed Budgeting (RAB) to support the smooth running of the project. In the process of running a ceiling work project, budget planning is used as the main source of information in calculating company profits. Details of the Budget The cost of a gypsum ceiling work project includes labor costs, experts, materials, project operations, work equipment costs and mobilization costs. Inappropriate planning can affect a company's cash flow so that it is closely related to the continuity of the implementation of ceiling work projects. A prediction

How to cite:

Kautsar Baihaqi Supriyadi, Feri Sulianta (tahun terbit) Analisis Prediksi Profit Proyek Pekerjaan Plafon Gypsum Menggunakan Klasifikasi Decision Tree, (4) 12, <https://doi.org/10.36418/syntax-idea.v3i6.1227>

E-ISSN:

2684-883X

Published by:

[Ridwan Institute](#)

technique is needed to determine company profits based on the use of a project's RAB so that these companies can maintain their cash flow so that it does not affect project implementation. The aim is to find out which factors are more decisive for generating profits in a ceiling work project. This study presents the implementation of data mining with classification methods, especially decision trees that can solve these problems and generate knowledge discovery to help companies gain future profits. The results of classification techniques can help companies pay more attention to certain attributes in implementing a ceiling work project and help companies or policy makers in determining a decision in the future.

Keywords: *Data mining; Klasifikasi; Decision tree; Prediction; Real Estimate of Cost; Profit*

Pendahuluan

Teknologi dan informasi mempunyai dampak yang sangat luas untuk berbagai bidang. Salah satunya yaitu bidang konstruksi (Abdillah et al., 2020). Dalam menjalankan kegiatan konstruksi, faktor utama yang menentukan keberhasilan suatu proyek adalah perencanaan anggaran (Rawis, Tjakra, & Arsjad, 2016). Dalam menentukan perencanaan anggaran, teknologi dan informasi dapat memberikan sebuah prediksi berapa banyak anggaran yang diperlukan dan meminimalisir tingkat kerugian (Purnaya & SE, 2016).

Dalam proyek pekerjaan plafon, perencanaan anggaran digunakan sebagai sumber informasi utama dalam menghitung keuntungan perusahaan. Perencanaan anggaran atau biasa disebut Rincian Anggaran Biaya (RAB), biasanya meliputi tenaga kerja, material, operasional proyek dan operasional perusahaan. Perencanaan yang tidak tepat dapat mempengaruhi cash flow sebuah perusahaan sehingga merupakan faktor utama dalam menentukan hasil suatu proyek pekerjaan plafon (Abma, Nugraheni, & Metalindra, 2020).

Berdasarkan permasalahan tersebut diperlukan adanya suatu teknik prediksi untuk menentukan keuntungan perusahaan berdasarkan penggunaan RAB suatu proyek. Teknik prediksi yang akan digunakan yaitu teknik klasifikasi data mining. Klasifikasi merupakan sebuah teknik data mining yang bertujuan untuk membuat sebuah model prediksi dengan memberikan label/kelas (Maharani, 2015). Dalam hal ini keuntungan perusahaan akan diberikan label yaitu “profit” atau “defisit”.

Dalam beberapa penelitian terdahulu analisis prediksi dengan menggunakan klasifikasi decision tree sudah diimplementasikan dalam berbagai bidang. Dalam penelitian (Wang & Gao, 2021);(Songze & Yan, 2005) teknik ini digunakan dalam bidang pendidikan untuk mengetahui korelasi antara kurikulum pendidikan terhadap prestasi siswa dan pemilihan karir setelah siswa tersebut lulus. Sedangkan dalam penelitian (Liu, Hu, Yan, & Zhang, 2013) metode ini digunakan untuk mendapatkan knowledge discovery terhadap performa karyawan yang sudah lulus dengan karyawan yang masih menempuh bidang pendidikan.

Dalam penelitian ini digunakan juga metode klasifikasi decision tree, yang diuji menggunakan data dalam bidang konstruksi khususnya pekerjaan plafon dengan

memprediksi keuntungan dan kerugian proyek tersebut berdasarkan beberapa faktor dan bertujuan untuk mencari faktor mana yang lebih menentukan untuk menghasilkan keuntungan dalam suatu proyek pekerjaan plafon.

Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan sistem yang digunakan untuk menemukan informasi dari database yang tersedia (Utami, 2019). Database berisi tabel-tabel yang mempunyai hubungan satu sama lain. Informasi tersebut ditransformasikan menjadi sebuah basis pengetahuan untuk dijadikan sebagai pertimbangan dalam pengambilan keputusan (Suyanto, 2017).

Knowledge Discovery in Database (KDD) dan data mining dikolaborasikan pada kondisi tertentu guna menjabarkan operasi penambangan informasi potensial yang belum ditemukan dalam sebuah database yang cukup besar (Faradillah, 2013). Dua hal tersebut secara konsep bisa disebut berbeda akan tetapi mempunyai kolerasi satu sama lain, bahkan salah satu tahapan dari prosedur KDD adalah data mining (Harahap, 2015). Berikut adalah tahapan dalam proses KDD:

1. Data Selection
2. Pre-processing / Cleaning
3. Transformation
4. Data mining
5. Evaluation

Data mining

Data mining adalah proses analisis terhadap knowledge discovery in database yang disingkat KDD. Pengetahuan bisa berupa pola (pattern) data atau relasi antar data yang valid (tidak diketahui sebelumnya) (Fayyad, Piatetsky-Shapiro, & Smyth, 1996);(Suyanto, 2017).

Data mining adalah gabungan berbagai ilmu komputer yang merupakan irisan dari artificial intelligence, machine learning, statistics, dan database systems (Geng & Hamilton, 2006);(Clifton, Niehaus, Charlton, & Colopy, 2015). Data mining memiliki dua fungsi yaitu deskriptif dan prediktif. Deskriptif artinya data mining digunakan untuk menemukan pola atau pattern yang menjelaskan karakteristik data. Terdapat 3 teknik data mining yang termasuk kedalam kategori deskriptif yaitu:

1. Klasterisasi (clustering)
2. Aturan asosiasi (association rule)
3. Sequential Pattern

Fungsi prediktif yaitu data mining digunakan dalam membentuk sebuah model pengetahuan yang akan digunakan untuk melakukan prediksi. Yang termasuk kedalam kategori ini yaitu:

1. Klasifikasi (classification)
2. Regresi (regression)
3. Deteksi anomaly (anomaly detection)

Klasifikasi

Klasifikasi adalah satu diantara teknik prediktif data mining. Klasifikasi dapat didefinisikan sebagai proses untuk menyatakan suatu objek data sebagai salah satu kategori yang telah didefinisikan sebelumnya (Zaki, Sayed, & Mori, 2013). Terdapat 4 teknik klasifikasi, diantaranya adalah:

1. Naïve Bayes
2. Decision Tree
3. ANN
4. SVM

Decision Tree

Decision tree disebut juga atau diartikan sebagai pohon keputusan. Metode decision tree mengubah informasi/fakta yang sangat besar menjadi sebuah pohon keputusan yang merpresentasikan aturan dan dapat diekspresikan dalam bentuk bahasa basis data seperti Structure Query Language (Anggraeni, Saputra, & Noranita, 2014). Pohon keputusan digunakan untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan variabel target (Achmad, Slamet, & Itats, 2012). Teknik klasifikasi yang menggunakan decision tree, biasanya berkaitan dengan Entropy dan Information Gain. Entropy merupakan sebuah parameter yang digunakan dalam mengukur keberagaman dalam suatu himpunan data (Pal & Pal, 1991). Secara matematis Entropy dirumuskan sebagai berikut:

$$Entropy(S) = \sum_{i=1}^c -p_i \log_2 p_i \quad (1)$$

Keterangan:

c : jumlah kelas

pi: perbandingan antara jumlah sampel di kelas i dengan jumlah sampel pada himpunan data

Information Gain adalah parameter efektivitas sebuah atribut dalam mengklasifikasikan data (Suyanto, 2017). Secara matematis dapat ditulis sebagai berikut:

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \left| \frac{S_i}{S} \right| Entropy(S_i) \quad (2)$$

Keterangan:

S: himpunan kasus

A: atribut

n : jumlah partisi atribut A

|Si|: jumlah sampel untuk nilai i

|S|: jumlah seluruh sampel data

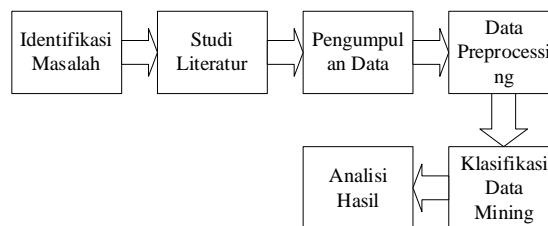
Profit/Keuntungan

Secara sederhana profit dapat diartikan sebagai selisih positif antara pendapatan dengan pengeluaran. Lebih sederhananya lagi profit yaitu pendapatan lebih besar dari pengeluaran (Barker, 2010). Profit merupakan meningkatnya nilai modal akibat adanya transaksi keluar masuknya uang yang memberikan pengaruh pada perusahaan dalam jangka waktu tertentu (Elisa, 2018).

Metode Penelitian

A. Desain Penelitian

Merupakan sebuah gambaran yang menyajikan setiap proses yang dilakukan dalam melaksanakan penelitian.



Gambar 1
Desain Penelitian

Berikut adalah tahapan-tahapan penelitian yang dilakukan sesuai dengan gambar diatas:

1. Identifikasi Masalah, proses pengidentifikasian permasalahan yang timbul sehingga diperlukan adanya penelitian.
2. Studi Literatur, proses pencarian informasi dari berbagai sumber yang berhubungan erat dengan permasalahan dalam penelitian.
3. Pengumpulan Data, pemilahan data yang dapat dikeluarkan oleh perusahaan, dikumpulkan menjadi satu sehingga menjadi *RAW* data yang akan diuji.
4. *Data Preprocessing*, proses sebelum data diolah menggunakan teknik klasifikasi.
5. Klasifikasi *Data mining*, proses pengklasifikasian *dataset*.
6. Analisis Hasil

B. Populasi dan Sampel

1. Populasi

Populasi pada penelitian ini adalah data *history* laporan proyek yang sudah selesai yang berisi persentasi penggunaan rincian anggaran biaya (RAB) terhadap nilai kontrak sebuah proyek di PT. Mitrakarya Teguhsetia pada periode 2018-2021.

2. Sampel

Sampel yang digunakan adalah 35 data sampel dari populasi data *history* laporan proyek yang sudah selesai di PT. Mitrakarya Teguhsetia periode 2018-2021.

C. Variabel Penelitian

Variabel merupakan sebuah objek yang mempengaruhi hasil penelitian. Dalam memprediksi *profit* terdapat beberapa variabel yang mempunyai korelasi terhadap *profit* itu sendiri, yaitu:

1. Nilai kontrak merupakan nilai pekerjaan yang dikerjakan sesuai dengan perjanjian atau Surat Perintah Kerja (SPK)
2. Waktu merupakan lama pelaksanaan proyek dari awal hingga proyek selesai.
3. Rincian Anggaran Biaya Tenaga Ahli (Upah) adalah biaya yang berupa pembayaran untuk upah tenaga ahli sebagai imbal jasa melakukan pengawasan terhadap pemasangan pekerjaan plafon dan partisi.
4. Rincian Anggaran Biaya Tenaga Kerja (Upah) adalah upah tenaga kerja sebagai imbal jasa dari jasa pemasangan pekerjaan plafon dan partisi.
5. Rincian Anggaran Biaya Material adalah biaya yang berupa pembayaran untuk bahan material yang akan digunakan dalam mengerjakan sebuah proyek.
6. Rincian Anggaran Biaya Operasional Proyek adalah pengeluaran perusahaan dalam menjalankan sebuah proyek.
7. Rincian Anggaran Biaya Peralatan adalah biaya yang berupa pengeluaran uang untuk peralatan yang digunakan pada saat proyek berjalan.
8. Rincian Anggaran Biaya Mobilisasi adalah biaya yang berupa pengeluaran uang untuk transportasi.

Hasil dan Pembahasan

A. Dataset

Penelitian ini menggunakan data rekap proyek pekerjaan plafon periode 2018-2021 sebagai dataset.

Tabel 1.
Dataset

PERSENTASE PERBANDINGAN PENGELUARAN/PENDAPATAN TERHADAP NILAI KONTRAK KESELURUHAN									
No	Nama Nilai ProyekKontrak	Waktu	Tenaga Ahli	Tenaga Kerja	Material	Operasion	Peralatan	Mobilisasi	Profit
1	PG001 11.252.92225	15%	26%	38%	8%	10%	10%	-7%	
2	PG002 84.556.23923	13%	19%	36%	9%	14%	11%	-2%	
3	PG003 46.245.22346	22%	33%	38%	6%	7%	7%	-13%	
4	PG004 48,648,17335	13%	21%	31%	1%	7%	7%	20%	
5	PG005 1.679.625. 314	6%	36%	41%	10%	8%	2%	-3%	
6	PG006 711,870,04272	22%	32%	39%	8%	12%	10%	-23%	
7	PG007 1.182.452. 542	21%	9%	32%	6%	8%	8%	16%	
8	PG008 518.922.74225	21%	24%	41%	1%	8%	11%	-6%	
9	PG009 87,029,04232	3%	9%	18%	12%	9%	15%	34%	
10	PG010 382,969,75256	11%	12%	32%	9%	12%	8%	16%	
11	PG011 526.235.94432	22%	32%	37%	7%	8%	9%	-15%	
12	PG012 113,380,8845	12%	22%	33%	6%	8%	11%	8%	

Analisis Prediksi Profit Proyek Pekerjaan Plafon Gypsum Menggunakan Klasifikasi
Decision Tree

13	PG0134,867,785, 926	25%	35%	36%	6%	9%	8%	-19%
14	PG014 862,948,6872	11%	8%	41%	9%	8%	13%	10%
15	PG015 27,035,25012	1%	7%	34%	5%	5%	10%	38%
16	PG0162.762.341. 330	21%	31%	18%	2%	8%	12%	8%
17	PG017 414,329,58142	8%	22%	15%	3%	13%	15%	24%
18	PG0182,558,848, 316	23%	33%	38%	10%	8%	8%	-20%
19	PG019 357,042,7992	3%	10%	36%	2%	8%	12%	29%
20	PG020 327,187,80314	24%	22%	15%	3%	9%	12%	15%
21	PG021 274,514,69122	7%	13%	37%	2%	8%	14%	19%
22	PG022 160,000,00138	11%	21%	37%	1%	8%	11%	11%
23	PG023 1,284,057, 183	16%	36%	36%	6%	9%	13%	-16%
24	PG024 35,615,36557	1%	16%	20%	6%	9%	10%	38%
25	PG025 56,458,47591	2%	18%	32%	6%	7%	7%	28%
26	PG026 171.828.2532	1%	18%	28%	1%	2%	10%	40%
27	PG027 95,884,40252	12%	34%	41%	8%	10%	11%	-16%
28	PG028 364,887,7287	9%	27%	38%	2%	13%	12%	-1%
29	PG029 101,819,9972	21%	33%	27%	12%	8%	1%	-2%
30	PG030 324.132.2337	1%	18%	33%	2%	10%	2%	34%
31	PG031 2.039.208. 186	8%	20%	36%	1%	12%	11%	12%
32	PG032 85,217,03331	1%	20%	36%	1%	8%	8%	26%
33	PG033 11.517.51220	11%	22%	21%	7%	11%	9%	19%
34	PG034 13.010.30014	1%	25%	10%	6%	10%	8%	10%
35	PG035 64.269.200148	6%	25%	38%	4%	2%	8%	17%

B. *Dat24a Preprocessing*

Berdasarkan kebutuhan dilakukan *data transformation* terhadap *dataset* yang ada pada sub-bab sebelumnya. Tipe data numerik akan ditransformasikan menjadi tipe data kategorical. Untuk variabel “nilai kontrak” akan ditransformasi menjadi 3 label dengan aturan sebagai berikut:

1. Nilai kontrak $< 250.000.000$ maka dilabeli *Low*
2. $250.000.000 \leq$ nilai kontrak $\leq 500.000.000$ dilabeli *Medium*
3. Nilai kontrak $> 500.000.000$ dilabeli *High*

Selanjutnya untuk variabel “waktu” akan dikategorikan menjadi 2 label yaitu “*Long*” dan “*Short*”. *Long* untuk jangka pengerjaan proyek >180 hari, sedangkan *short* untuk proyek yang pengerjaannya <180 hari. (Faradillah, 2013) Untuk 6 variabel lainnya yaitu Tenaga Ahli, Tenaga Kerja, Material, Operasional Proyek, Peralatan dan Mobilisasi akan dikategorikan menjadi 3 label yaitu *Underspend* (pengeluaran dibawah *budget RAB*), *Budget* (pengeluaran sesuai dengan *budget RAB*) dan *Overspend* (pengeluaran diatas *budget RAB*). Aturan yang ditetapkan untuk ke-6 variabel tersebut berbeda-beda tergantung dengan persentase masing masing variable terhadap RAB dengan rincian sebagai berikut:

Tabel 2
Transformasi Tenaga Ahli

LABEL	RAB (X)
<i>Underspend</i>	$\leq 10\%$
<i>Budget</i>	$20\% \geq X > 10\%$
<i>Overspend</i>	$> 20\%$

Tabel 3
Tranformasi Tenaga Kerja

LABEL	RAB (X)
<i>Underspend</i>	$\leq 20\%$
<i>Budget</i>	$30\% \geq X > 20\%$
<i>Overspend</i>	$> 30\%$

Tabel 4
Transformasi Material

LABEL	RAB (X)
<i>Underspend</i>	$\leq 30\%$
<i>Budget</i>	$35\% \geq X > 30\%$
<i>Overspend</i>	$> 35\%$

Tabel 5
Transformasi Operasional Proyek

LABEL	RAB (X)
<i>Underspend</i>	$\leq 5\%$
<i>Budget</i>	$8\% \geq X > 5\%$
<i>Overspend</i>	$> 8\%$

Tabel 6
Transformasi Peralatan

LABEL	RAB (X)
<i>Underspend</i>	$\leq 7\%$
<i>Budget</i>	$10\% \geq X > 7\%$
<i>Overspend</i>	$> 10\%$

Tabel 7
Transformasi Mobilisasi

LABEL	RAB (X)
<i>Underspend</i>	$\leq 7\%$
<i>Budget</i>	$10\% \geq X > 7\%$
<i>Overspend</i>	$> 10\%$

Sedangkan untuk profit/keuntungan akan dikategorikan menjadi 2 kelas yaitu Profit bila keuntungan didapat $\geq 10\%$ dan Defisit bila keuntungan yang didapat $< 10\%$ atau bahkan negatif. Hasil akhir dari rangkaian proses *Data Preprocessing* adalah sebagai berikut:

Tabel 8
Format Dataset setelah Preprocessing

DATASET SETELAH MELALUI DATA PREPROCESSING										
No	Nama Proyek	Nilai Kontrak	Waktu	Tenaga Ahli	Tenaga Kerja	Material	Operasional Proyek	Peralatan	Mobilisasi	Profit
1	PG001	Low	Short	Budget	Budget	Overspend	Overspend	Budget	Overspend	Defisit
2	PG002	Low	Short	Budget	Underspend	Overspend	Budget	Budget	Overspend	Defisit
3	PG003	Low	Short	Overspend	Overspend	Overspend	Budget	Budget	Budget	Defisit
4	PG004	Low	Short	Budget	Budget	Budget	Underspend	Budget	Budget	Profit
5	PG005	High	Long	Underspend	Overspend	Overspend	Overspend	Budget	Underspend	Defisit
6	PG006	High	Long	Overspend	Overspend	Overspend	Budget	Overspend	Budget	Defisit
7	PG007	High	Long	Overspend	Underspend	Budget	Budget	Underspend	Budget	Profit
8	PG008	High	Long	Overspend	Budget	Overspend	Underspend	Budget	Overspend	Defisit
9	PG009	Low	Short	Underspend	Underspend	Budget	Overspend	Budget	Overspend	Profit
10	PG010	Medium	Long	Budget	Underspend	Budget	Overspend	Overspend	Budget	Profit
11	PG011	High	Long	Overspend	Overspend	Overspend	Budget	Budget	Budget	Defisit
12	PG012	Low	Short	Budget	Budget	Budget	Budget	Budget	Overspend	Profit
13	PG013	High	Long	Overspend	Overspend	Overspend	Budget	Budget	Budget	Defisit
14	PG014	High	Short	Budget	Underspend	Overspend	Overspend	Budget	Overspend	Profit
15	PG015	Low	Short	Underspend	Underspend	Budget	Underspend	Underspend	Budget	Profit
16	PG016	High	Long	Overspend	Overspend	Underspend	Underspend	Budget	Overspend	Profit
17	PG017	Medium	Short	Underspend	Budget	Underspend	Underspend	Overspend	Overspend	Profit
18	PG018	High	Long	Overspend	Overspend	Overspend	Overspend	Budget	Budget	Defisit
19	PG019	Medium	Short	Underspend	Underspend	Overspend	Underspend	Budget	Overspend	Profit
20	PG020	Medium	Long	Overspend	Budget	Underspend	Underspend	Budget	Overspend	Profit
21	PG021	Medium	Short	Underspend	Underspend	Overspend	Underspend	Budget	Overspend	Profit
22	PG022	Medium	Short	Budget	Budget	Overspend	Underspend	Budget	Overspend	Profit
23	PG023	High	Long	Budget	Overspend	Overspend	Budget	Budget	Overspend	Defisit
24	PG024	Low	Short	Underspend	Underspend	Underspend	Budget	Budget	Budget	Profit
25	PG025	Low	Short	Underspend	Underspend	Budget	Budget	Underspend	Underspend	Profit
26	PG026	Medium	Short	Underspend	Underspend	Underspend	Underspend	Underspend	Budget	Profit
27	PG027	Medium	Short	Budget	Overspend	Overspend	Budget	Budget	Overspend	Defisit
28	PG028	Medium	Short	Underspend	Budget	Overspend	Underspend	Overspend	Overspend	Defisit
29	PG029	Medium	Short	Overspend	Overspend	Underspend	Overspend	Budget	Underspend	Defisit
30	PG030	High	Short	Underspend	Underspend	Budget	Underspend	Budget	Underspend	Profit
31	PG031	High	Long	Budget	Underspend	Overspend	Underspend	Overspend	Overspend	Profit
32	PG032	Low	Short	Underspend	Underspend	Overspend	Underspend	Budget	Budget	Profit
33	PG033	Low	Short	Budget	Budget	Underspend	Budget	Overspend	Budget	Profit
34	PG034	Low	Short	Underspend	Budget	Overspend	Budget	Budget	Budget	Profit
35	PG035	Low	Short	Underspend	Budget	Overspend	Underspend	Underspend	Budget	Profit

C. Klasifikasi

Dari Data Preprocessing yang telah dihasilkan, selanjutnya dilakukan proses klasifikasi dengan menggunakan metode Decision tree. Dari tabel tersebut akan ditentukan atribut/variabel mana yang menjadi akar atau node, berdasarkan nilai Information Gain tertinggi yang dimiliki oleh masing-masing atribut. Dalam menentukan nilai Information Gain diperlukan nilai Entropy masing-masing atribut terlebih dahulu. Nilai Entropy dapat dihitung dengan menggunakan rumus(1). Selanjutnya menghitung Information Gain dengan rumus(2).

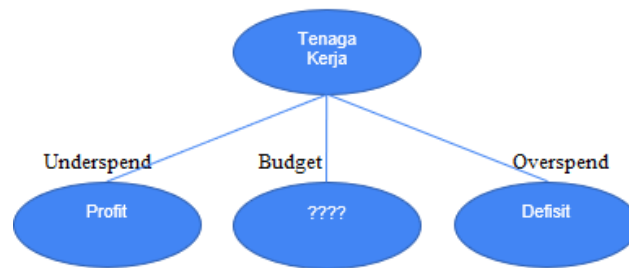
Perhitungan lengkap untuk nilai *Entropy* dan *Information Gain* masing-masing atribut yang digunakan disajikan dalam bentuk tabel sebagai berikut:

Tabel 9
Perhitungan Node 1

Node	Atribut	Nilai	Jumlah Kasus	Profit	Defisit	Entropy	Gain
1	Total		35	22	13	0,952	

Nilai Kontrak					0,074
	<i>Low</i>	13	10	3	0,779
	<i>Medium</i>	10	7	3	0,881
	<i>High</i>	12	5	7	0,980
Waktu					0,071
	<i>Short</i>	23	17	6	0,828
	<i>Long</i>	12	5	7	0,980
Tenaga Ahli					0,166
	<i>Underspend</i>	14	12	2	0,592
	<i>Budget</i>	11	7	4	0,946
	<i>Overspend</i>	10	3	7	0,881
Tenaga Kerja					0,404
	<i>Underspend</i>	14	13	1	0,371
	<i>Budget</i>	11	8	3	0,845
	<i>Overspend</i>	10	1	9	0,469
Material					0,279
	<i>Underspend</i>	7	6	1	0,592
	<i>Budget</i>	8	8	0	0
	<i>Overspend</i>	20	8	12	0,971
Operasional Proyek					0,142
	<i>Underspend</i>	15	13	2	0,567
	<i>Budget</i>	13	6	7	0,996
	<i>Overspend</i>	7	3	4	0,985
Peralatan					0,112
	<i>Underspend</i>	5	5	0	0
	<i>Budget</i>	24	13	11	0,994
	<i>Overspend</i>	6	4	2	0,918
Mobilisasi					0,007
	<i>Underspend</i>	4	2	2	1
	<i>Budget</i>	15	10	5	0,918
	<i>Overspend</i>	16	10	6	0,954

Dari perhitungan tabel diketahui bahwa tenaga kerja merupakan nilai dengan gain terbesar yaitu 0,404 maka Tenaga Kerja dijadikan sebuah akar/node. Dari nilai yang terdapat dalam tabel perhitungan tersebut kita dapat menyimpulkan ketika sebuah proyek pekerjaan plafon biaya tenaga kerjanya underspend dapat dipastikan proyek tersebut akan mendapatkan profit. Sedangkan ketika tenaga kerja overspend proyek tersebut sudah dipastikan akan defisit. Ketika tenaga kerja yang dikeluarkan sesuai budget masih belum dapat dilihat hasil keuntungan proyeknya. Maka dari itu pohon keputusan yang terbentuk juga belum terisi secara penuh.



Gambar 2
Pohon Keputusan Node 1

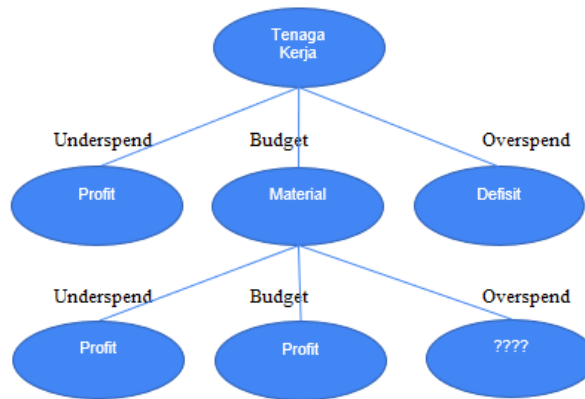
Oleh karena itu kita perlu melakukan perhitungan kembali. Perhitungan ini disebut node 1.1. Akan tetapi perhitungan node 1.1 tidak dilakukan terhadap semua data seperti node 1. Perhitungan dilakukan terhadap data yang telah difilter dengan ketentuan nilai atribut Tenaga Kerja adalah “Budget” dan dihasilkan data sebagai berikut:

Tabel 10
Perhitungan Node 1.1

Node	Atribut	Nilai	Jumlah Kasus	Profit	Defisit	Entropy	Gain
1.1	Total (Tenaga Kerja [Budget])		11	8	3	0,845	
	Nilai Kontrak						0,196
		<i>Low</i>	6	5	1	0,650	
		<i>Medium</i>	4	3	1	0,811	
		<i>High</i>	1	0	1	0	
	Waktu						0,220
		<i>Short</i>	9	7	2	0,764	
		<i>Long</i>	2	1	1	0	
	Tenaga Ahli						0,040
		<i>Underspend</i>	4	3	1	0,811	
		<i>Budget</i>	5	4	1	0,722	
		<i>Overspend</i>	2	1	1	1	
	Material						0,300
		<i>Underspend</i>	3	3	0	0	
		<i>Budget</i>	2	2	0	0	
		<i>Overspend</i>	6	3	3	1	
	Operasional Proyek						0,296
		<i>Underspend</i>	7	5	2	0,863	
		<i>Budget</i>	3	3	0	0	
		<i>Overspend</i>	1	0	1	0	
	Peralatan						0,045
		<i>Underspend</i>	1	1	0	0	
		<i>Budget</i>	7	5	2	0,863	

	<i>Overspend</i>	1	0	1	0
Mobilisasi					0,218
	<i>Underspend</i>	0	0	0	0
	<i>Budget</i>	4	4	0	0
	<i>Overspend</i>	7	4	3	0,985

Berdasarkan perhitungan tersebut, didapatkan hasil bahwa Material merupakan akar untuk *node 1.1* (*Information Gain* tertinggi). Maka pohon keputusannya adalah:



Gambar 3
Pohon Keputusan Node 1.1

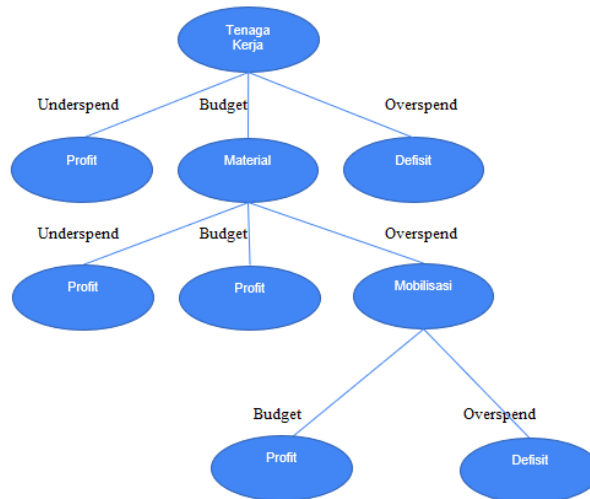
Dalam *Node 1.1* ketika *Material Overspend* belum dapat diketahui keuntungan yang didapat oleh perusahaan apakah *profit* atau *defisit* sehingga perlu dilakukan kembali perhitungan selanjutnya terhadap data *node 1.1* dengan nilai atribut *Material* adalah “*Overspend*” dan akan disebut dengan *Node 1.2*.

Tabel 11
Perhitungan Node 1.2

Node	Atribut	Nilai	Jumlah Kasus	<i>Profit</i>	<i>Defisit</i>	<i>Entropy</i>	Gain
1.2	Total (Material [Overspend])		6	3	3	1	
	Nilai Kontrak						0,208
		<i>Low</i>	3	2	1	0,918	
		<i>Medium</i>	2	1	1	1	
		<i>High</i>	1	0	1	0	
	Waktu						0,191
		<i>Short</i>	5	3	2	0,971	
		<i>Long</i>	1	0	1	0	
	Tenaga Ahli						0,208
		<i>Underspend</i>	3	2	1	0,918	
		<i>Budget</i>	2	1	1	1	
		<i>Overspend</i>	1	0	1	0	
	Operasional Proyek						0,333

	<i>Underspend</i>	4	2	2	1
	<i>Budget</i>	1	0	0	0
	<i>Overspend</i>	1	0	1	0
Peralatan					0,333
	<i>Underspend</i>	1	1	0	0
	<i>Budget</i>	4	2	2	1
	<i>Overspend</i>	1	0	1	0
Mobilisasi					0,459
	<i>Underspend</i>	0	0	0	0
	<i>Budget</i>	2	2	0	0
	<i>Overspend</i>	4	1	3	0,811

Dalam *node* 1.2 atribut dengan nilai *Information Gain* tertinggi adalah Mobilisasi, dengan hasil pohon keputusan sebagai berikut.



Gambar 4
Pohon Keputusan Node 1.2

Perhitungan dihentikan sampai dengan *node* 1.2 karena pohon keputusan sudah terbentuk dan semua akarnya sudah terisi.

D. Analisis dan Hasil

Pohon keputusan tersebut menghasilkan aturan/*rule* berdasarkan *knowledge discovery* yang didapat sebagai berikut:

1. Jika biaya “Tenaga Kerja” *overspend* maka proyek pekerjaan plafon akan mengalami kerugian (*defisit*).
2. Jika biaya “Tenaga Kerja” *underspend* maka proyek akan mengalami keuntungan (*profit*).
3. Jika biaya “Tenaga Kerja” *budget* dan biaya “Material” *underspend* atau *budget* maka proyek pekerjaan plafon akan mengalami keuntungan (*profit*).
4. Jika biaya “Tenaga Kerja” *budget* dan biaya “material” *overspend*, biaya “Mobilisasi” menjadi penentu. Jika *budget* perusahaan mengalami *profit* sedangkan jika *overspend* perusahaan mengalami *defisit*.

Berdasarkan *knowledge discovery* yang dihasilkan, terdapat beberapa *rules* yang sama dengan yang dialami oleh perusahaan, yaitu:

1. Biaya “Tenaga Kerja” dan “Material” yang sering melebihi anggaran merupakan penyebab faktor utama perusahaan mengalami kerugian dalam sebuah proyek.
2. Dalam beberapa proyek, ketika perusahaan mampu menekan biaya “Tenaga Kerja” dan “Material”, proyek tersebut selalu mendapatkan keuntungan.

Kesimpulan

Metode klasifikasi Decision tree dapat digunakan dalam memprediksi profit proyek pekerjaan plafon di PT. Mitrakarya Teguhsetia dengan atribut yang terpenting yaitu RAB Tenaga Kerja. Hal tersebut sesuai dengan keadaan dalam perusahaan dimana dalam setiap proyek yang berjalan, dimana jika RAB Tenaga Kerja selalu memegang peranan penting dalam setiap proyeknya.

Pohon keputusan yang dihasilkan oleh teknik klasifikasi dapat membantu perusahaan untuk lebih memperhatikan atribut tertentu dalam pelaksanaan sebuah proyek pekerjaan plafon dan membantu perusahaan atau pemangku kebijakan dalam menentukan sebuah keputusan di masa depan.

Bibliografi

- Abdillah, Leon A., Alwi, Moh Hatta, Simarmata, Janner, Bisyrri, Muhammad, Nasrullah, Nasrullah, Asmeati, Asmeati, Gusty, Sri, Sakir, Sakir, Affandy, Nur Azizah, & Bachtiar, Erniati. (2020). *Aplikasi Teknologi Informasi: Konsep dan Penerapan*. Yayasan Kita Menulis. [Google Scholar](#)
- Abma, Vendie, Nugraheni, Fitri, & Metalindra, Metalindra. (2020). Cash Flow Proyek Dengan Sumber Modal Bank Syariah Pada Pembangunan Dan Rehabilitasi Gedung Pelayanan Kesehatan Pemerintah Kabupaten Gunung Kidul. *Construction and Material Journal*, 2(2), 77–90. [Google Scholar](#)
- Achmad, Budanis Dwi Meilani, Slamet, Fauzi, & ITATS, Fakultas Teknologi Informasi. (2012). Klasifikasi data karyawan untuk menentukan jadwal kerja menggunakan metode decision tree. *Jurnal Iptek*, 16(1). [Google Scholar](#)
- Anggraeni, Hapsari Dita, Saputra, Ragil, & Noranita, Beta. (2014). Aplikasi Data Mining Analisis Data Transaksi Penjualan Obat Menggunakan Algoritma Apriori (Studi Kasus di Apotek Setya Sehat Semarang). *Journal of Informatics and Technology*, 2(2), 24–30. [Google Scholar](#)
- Barker, Richard. (2010). On the definitions of income, expenses and profit in IFRS. *Accounting in Europe*, 7(2), 147–158. [Google Scholar](#)
- Clifton, David A., Niehaus, K. E., Charlton, Peter, & Colopy, G. W. (2015). Health informatics via machine learning for the clinical management of patients. *Yearbook of Medical Informatics*, 24(01), 38–43. [Google Scholar](#)
- Elisa, Erlin. (2018). Prediksi Profit Pada Perusahaan Dengan Klasifikasi Algoritma C4. *5. Kumpulan Jurnal Ilmu Komputer (Klik)*, 5(02), 179–189. [Google Scholar](#)

- Faradillah, Sarah. (2013). Implementasi Data Mining Untuk Pengenalan Karakteristik Transaksi Customer Dengan Menggunakan Algoritma C4. 5. *Pelita Informatika Budi Darma*, 5(3), 1–5. [Google Scholar](#)
- Fayyad, Usama, Piatetsky-Shapiro, Gregory, & Smyth, Padhraic. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37. [Google Scholar](#)
- Geng, Liqiang, & Hamilton, Howard J. (2006). Interestingness measures for data mining: A survey. *ACM Computing Surveys (CSUR)*, 38(3), 9-es. [Google Scholar](#)
- Harahap, Fitriana. (2015). Penerapan Data Mining dalam Memprediksi Pembelian cat. *Proceedings Konferensi Nasional Sistem Dan Informatika (KNS&I)*. [Google Scholar](#)
- Liu, Yue, Hu, Lingjie, Yan, Fei, & Zhang, Bofeng. (2013). Information gain with weight based decision tree for the employment forecasting of undergraduates. *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, 2210–2213. IEEE. [Google Scholar](#)
- Maharani, Warih. (2015). Klasifikasi Data Menggunakan JST Backpropagation Momentum dengan Adaptive Learning Rate. *Seminar Nasional Informatika (SEMNASIF)*, 1(4). [Google Scholar](#)
- Pal, Nikhil R., & Pal, Sankar K. (1991). Entropy: A new definition and its applications. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(5), 1260–1270. [Google Scholar](#)
- Purnaya, I. Gusti Ketut, & SE, S. H. (2016). *Manajemen Sumber Daya Manusia*. Penerbit Andi. [Google Scholar](#)
- Rawis, Thresia Deisy, Tjakra, Jermias, & Arsjad, Tisano Tj. (2016). Perencanaan biaya Keselamatan dan Kesehatan Kerja (K3) pada proyek konstruksi bangunan (studi kasus: sekolah st. ursula kotamobagu). *Jurnal Sipil Statik*, 4(4). [Google Scholar](#)
- Songze, Lei, & Yan, Hao. (2005). Data mining in employment based on decision tree [J]. *Journal of Xi'an Institute of Technology*, 25(5), 429–432. [Google Scholar](#)
- Suyanto, Dr. (2017). Data Mining untuk klasifikasi dan klusterisasi data. *Bandung: Informatika Bandung*. [Google Scholar](#)
- Utami, Nur Rina. (2019). *Knowledge discovery in database metode generalized vector space model pada sistem temu kembali informasi terjemahan Kitab Mizanul Hikmah Berbahasa Indonesia*. Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah. [Google Scholar](#)

Kautsar Baihaqi Supriyadi, Feri Sulianta

Wang, Huan Bin, & Gao, Yang Jun. (2021). Research on C4. 5 algorithm improvement strategy based on MapReduce. *Procedia Computer Science*, 183, 160–165. [Google Scholar](#)

Zaki, Mohamed H., Sayed, Tarek, & Mori, Greg. (2013). Classifying road users in urban scenes using movement patterns. *Journal of Computing in Civil Engineering*, 27(4), 395–406. [Google Scholar](#)

Copyright holder:

Kautsar Baihaqi Supriyadi, Feri Sulianta (2022)

First publication right:

[Syntax Idea](#)

This article is licensed under:

