

```
/Applications/Anaconda-Navigator.app/Contents/MacOS/python /Users/
kycchung/Desktop/Capstone-master/evaluation_word_doc.py
```

Character accuracy in Tesseract/App_form

```
Characters predicted  7404.00
                   Errors    90.00
Character accuracy    98.78
```

Throughput in Tesseract/App_form

```
Characters predicted  7404.00
                   Errors    90.00
                   OCR time   16.24
                   Page      13.00
OCR time per page    1.25
```

	Penalty	Throughput
0	0	455.91
1	1	450.37
2	2	444.83
3	3	439.29
4	4	433.74
5	5	428.20
6	6	422.66
7	7	417.12
8	8	411.58
9	9	406.03

Accuracy by Character Class in Tesseract/App_form

```
Characters predicted  7404
Errors    90
```

	Class_type	Count_predicted	Count_error	Accuracy
0	Lowercase	6182	56	99.09
1	Uppercase	789	13	98.35
2	Number	160	8	95.00
3	Symbol	273	13	95.24

Accuracy by frequency in Tesseract/App_form

	Character	Count_predicted	Count_error	Accuracy_by_frequency
0	"	10	0	100.00
1	#	1	1	0.00
2	\$	2	0	100.00
3	%	1	0	100.00
4	'	5	0	100.00
5	(40	1	97.50
6)	40	0	100.00
7	*	1	0	100.00
8	+	1	1	0.00
9	,	32	0	100.00
10	/	20	0	100.00
11	0	15	0	100.00
12	1	28	1	96.43
13	2	40	1	97.50

14	3	18	3	83.33
15	4	12	2	83.33
16	5	17	1	94.12
17	6	7	0	100.00
18	7	11	0	100.00
19	8	8	0	100.00
20	9	4	0	100.00
21	:	115	10	91.30
22	?	1	0	100.00
23	@	4	0	100.00
24	A	76	0	100.00
25	B	13	0	100.00
26	C	67	3	95.52
27	D	33	2	93.94
28	E	55	0	100.00
29	F	30	0	100.00
30	G	10	0	100.00
31	H	22	0	100.00
32	I	51	5	90.20
33	K	16	0	100.00
34	L	24	0	100.00
35	M	31	0	100.00
36	N	81	1	98.77
37	O	44	0	100.00
38	P	52	0	100.00
39	Q	1	0	100.00
40	R	52	0	100.00
41	S	30	1	96.67
42	T	70	0	100.00
43	U	7	0	100.00
44	V	12	0	100.00
45	W	4	0	100.00
46	X	1	0	100.00
47	Y	7	1	85.71
48	a	554	6	98.92
49	b	89	0	100.00
50	c	226	3	98.67
51	d	224	3	98.66
52	e	724	5	99.31
53	f	161	2	98.76
54	g	77	1	98.70
55	h	198	1	99.49
56	i	528	4	99.24
57	j	1	0	100.00
58	k	20	0	100.00
59	l	311	2	99.36
60	m	186	2	98.92
61	n	456	4	99.12
62	o	571	6	98.95
63	p	214	2	99.07
64	q	8	0	100.00
65	r	395	0	100.00
66	s	410	5	98.78
67	t	510	5	99.02
68	u	142	1	99.30
69	v	43	0	100.00
70	w	40	0	100.00

71	x	21	0	100.00
72	y	67	4	94.03
73	z	6	0	100.00

Grouped error in Tesseract/App_form

Count_error

Character Predicted

#	*	1
(C	1
+	*	1
1	{}	1
2	{}	1
3	B	1
	{}	2
4	A	1
	{}	1
5	{}	1
:	2	5
	t	1
	{}	4
C	{}	3
D	{}	2
I	L	2
	T	1
	{}	2
N	{}	1
S	{}	1
Y	{}	1
a	A	2
	{}	4
c	C	1
	e	1
	{}	1
d	{}	3
e	{}	5
f	{}	2
g	{}	1
h	{}	1
i	{}	4
l	{}	2
m	M	1
	{}	1
n	N	2
	{}	2
o	O	2
	{}	4
p	P	1
	{}	1
s	{}	5
t	{}	5
u	{}	1
y	Y	2
	{}	2

Note: {} indicates no prediction made by the OCR engine
Number of missed prediction: 64

Error table with prediction confidence for Tesseract/App_form

	Character	Predicted	OCR confidence
0	:	2	0.0
1	:	t	0.0
2	:	2	0.0
3	I	T	11.0
4	:	2	35.0
5	:	2	38.0
6	:	2	45.0
7	I	L	57.0
8	o	O	59.0
9	#	*	67.0
10	n	N	72.0
11	a	A	72.0
12	y	Y	72.0
13	3	B	74.0
14	c	e	74.0
15	(C	76.0
16	4	A	86.0
17	n	N	90.0
18	a	A	90.0
19	p	P	90.0
20	m	M	90.0
21	o	O	90.0
22	c	C	90.0
23	y	Y	90.0
24	+	*	91.0
25	I	L	91.0

Average confidence: 63.46

Process finished with exit code 0