



KONICA MINOLTA

Document Functional Type Classification

Kristýna Němcová

Functional types

- Brochure
- Contract
- Financial Report
- Invoice
- Meeting minute, memo
- NDA
- Order (Purchase order)
- Patent
- Project charter (plan, gantt)
- Project status report
- Questionnaire
- Scientific article
- Technical Specification

Document samples

Arboriculture & Urban Forestry 36(5): September 2010

193



Arboriculture & Urban Forestry 2010. 36(5): 193-205



Tree Root Ecology in the Urban Environment and Implications for a Sustainable Rhizosphere

Susan D. Day, P. Eric Wiseman, Sarah B. Dickinson, and J. Roger Harris

Abstract. This review examines current understandings of how the belowground characteristics of urban settings affect tree roots as well as how tree roots contribute to biogeochemical processes in this belowground environment. Soil characteristics common to the urban environment include soil compaction and other physical impediments to root exploration, elevated pH, altered temperature and moisture patterns, and the presence of contaminants. These conditions may alter the growth dynamics, morphology, and physiology of roots. At the same time, roots have a profound effect on the soil environment, with trees directing 40%–73% of assimilated carbon below ground. Urban rhizosphere ecology is a topic of renewed interest for research not only because of its critical role in the urban ecosystem, but also because of its role in global environmental issues. In addition to its obvious contribution to aboveground growth, root exploration of the soil environment can influence environmental sustainability through root contributions to soil structure and drainage. Root influence is further mediated by the intimate role of roots in soil biological activity and thus carbon storage and nutrient cycling. Current advances and implications for emerging research are discussed.

Key Words. Heavy Metals; Road Salt; Root Periodicity; Soil Compaction; Soil Structure; Urban Hydrology; Urban Infrastructure.

INDEPENDENT CONTRACTOR AGREEMENT

This Independent Contractor Agreement (“Agreement”) is made effective as of the date set forth below by and between Commerce Technologies Corporation dba E-Commerce Exchange Los Angeles and IMS of CTC, located at 2321 Rosecrans Ave., Suite 4270, El Segundo, CA 90245 (“CTC”) and the entity and/or individual whose name and address are set forth below on the signature page for this Agreement (hereinafter referred to as “Independent Contractor”).

RECITALS

WHEREAS, Independent Contractor engages in the business of marketing services to business entities that accept credit cards as payment for goods and services;

WHEREAS, CTC wishes to expand its market share by retaining Independent Contractor to assist in marketing its credit card, debit card, EBT, gift card, loyalty card, leasing, ACH, POS equipment, software and related good and services (the “CTC Services”) and Independent Contractor wishes to undertake certain duties and responsibilities for marketing the CTC Services; and

Invoice



Sample Corporation
123 Corporate Lane
London, UK
Tel: +3 555-555-5555
Fax: +3 555-555-1234

Date: 4/19/2011

Invoice to: Cisco Systems
Address: 350 E Tasman Rd
San Jose, CA 95134

Invoice No. 98100072
PO No. 201492019

Cisco Contact
John Cisco
Tel: 555-555-5512

Vendor Contact
Nancy Vendor
Tel: +3 555-555-5534

Goods Sold

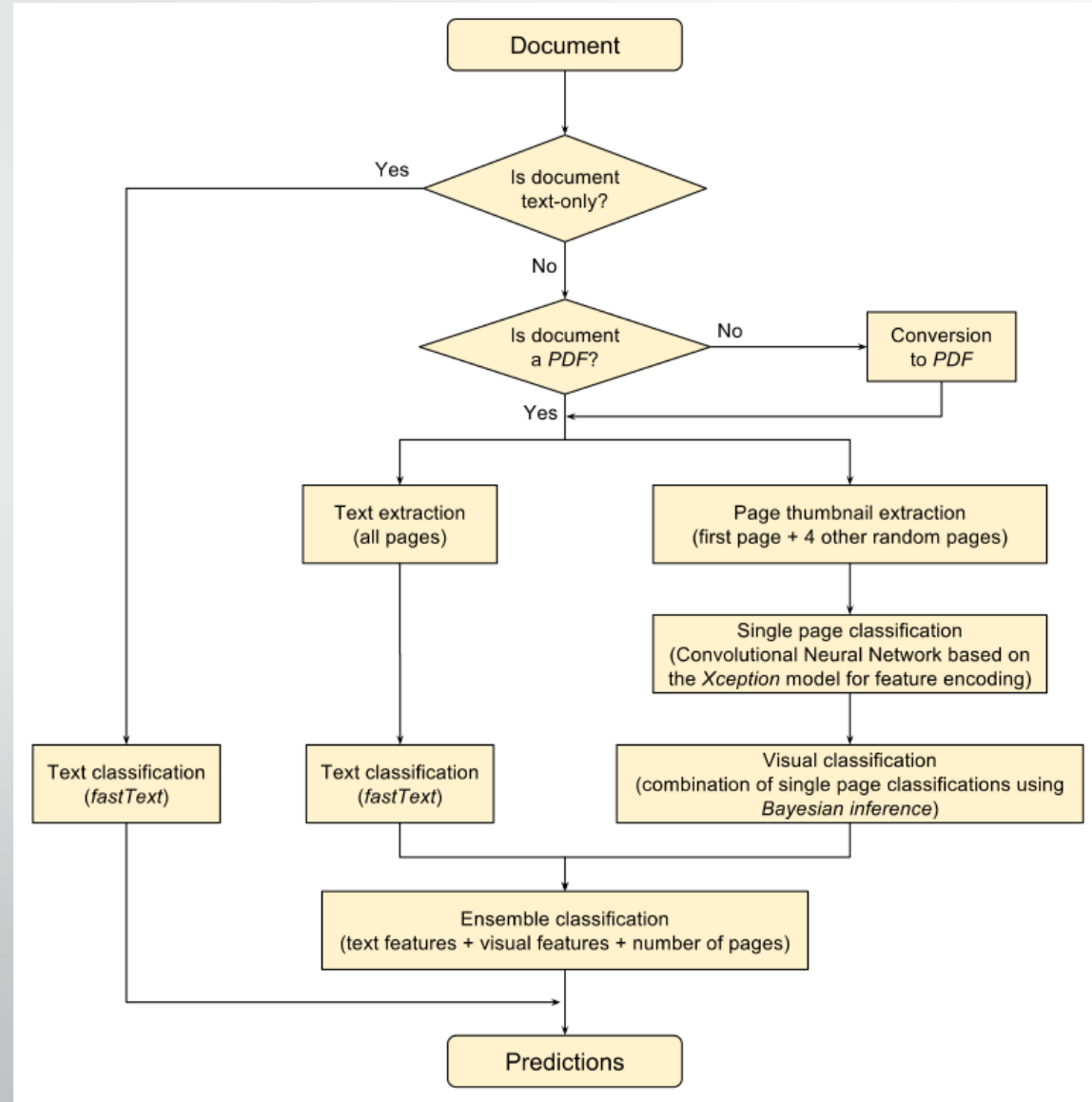
Part No.	Description	Quantity	Unit Price	Amount
32918	Server	6	2,000.00	12,000.00
82041	Widgets	40	100.00	4,000.00

Total Invoice Amount

16,000.00

Payments Made To:
Sample Corporation
123 Corporate Lane
London, UK

Flow chart



Preprocessing

- Documents to text
- Text to vectors
 - fastText pre-trained vectors (1 million word vectors trained on Wikipedia 2017)
- Documents to image (first twenty pages)

Text-based Classifier

- Two layer net
 - Input: average vector of text, size 300
 - Hidden layer: size 80
 - Output: functional types, size 13

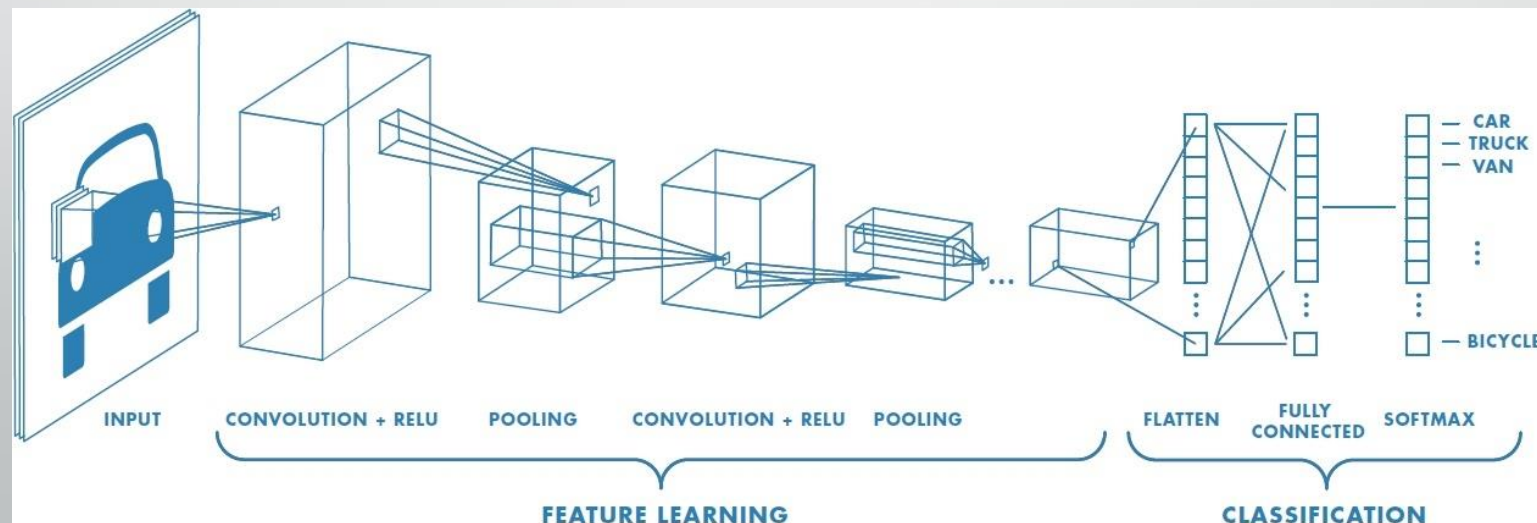
Layout-based Classifier

Single page classification

- Retraining pretrained ResNet-152
- L2 penalty to prevent over-fitting

Combine results of pages

- Average of the resulting vectors

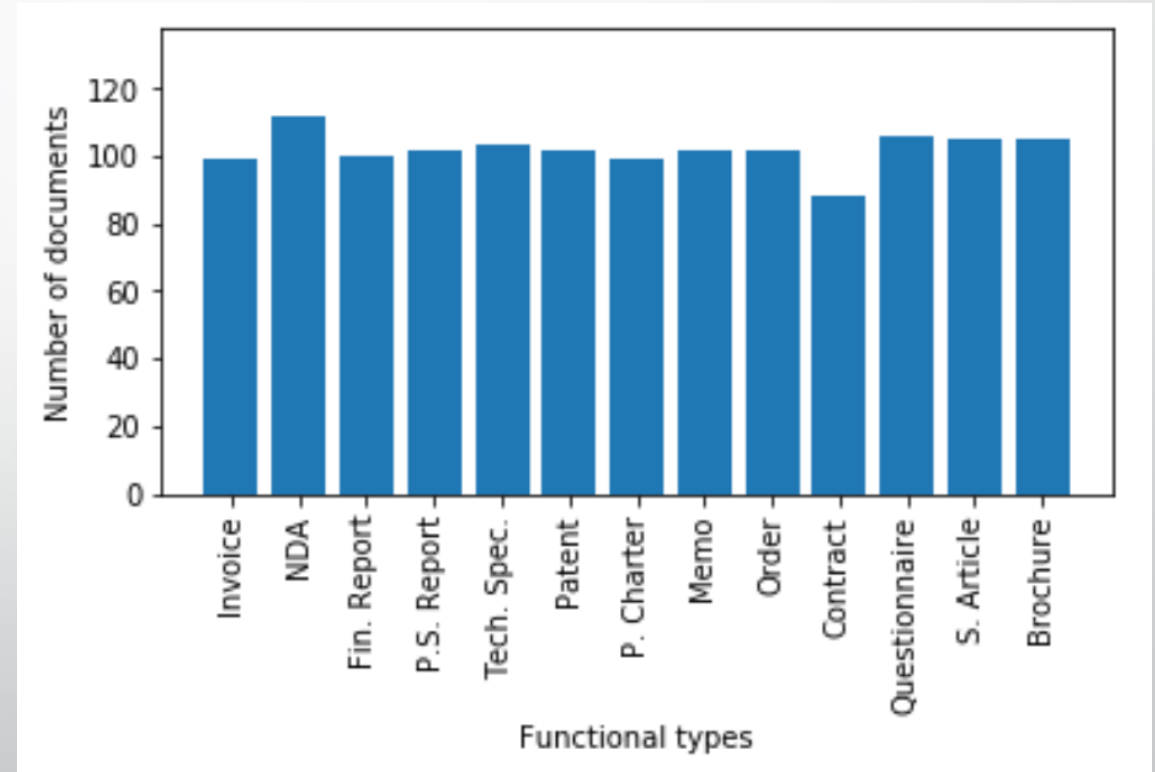


Final Classifier

- Two layer net
 - Input: set of confidences from text and layout part, size 26
 - Hidden layer: size 200
 - Output: functional types, size 13

Dataset

- Small dataset (approx 13 x 100)
- 80 % of documents from each category belongs to the training dataset
- 20 % test dataset
- For final prediction
 - 80 % from 20 % for training



Evaluation

- Text-based classifier:
 - 85,5 %
- Layout-based classifier :
 - Single page: 55 %
 - Whole document: 46,5 %
- Final:
 - 81 % (inconclusive)

Conclusion

- Text-based classification is the most important part in final classification .
- The layout-part provides more stable and robust final result.

Future work

- Whole training dataset for final classification
- Improve overall visual result of whole documents.
- Improve text vectorization of whole document.