BESCHWERDEKAMMERN   BOARDS OF APPEAL OF   CHAMBRES DE RECOURS
DES EUROPÄISCHEN   THE EUROPEAN PATENT   DE L'OFFICE EUROPÉEN
PATENTAMTS   OFFICE   DES BREVETS

Internal distribution code:
(A) [-] Publication in OJ
(B) [-] To Chairmen and Members
(C) [-] To Chairmen
(D) [X] No distribution

Datasheet for the decision
of 6 July 2022

Case Number: T 0874/19 - 3.5.07

Application Number: 14725807.3

Publication Number: 2973038

IPC: G06F17/30, G06K9/62, G06N3/04, G06N3/08

Language of the proceedings: EN

Title of invention:
CLASSIFYING RESOURCES USING A DEEP NETWORK

Applicant:
Google LLC

Headword:
Classifying resources using a deep network/GOOGLE

Relevant legal provisions:
EPC Art. 56

Keyword:
Inventive step - main and first to second auxiliary requests
(no)

Decisions cited:
T 0872/19, T 1227/05, G 0001/19

Case Number: **T 0874/19 - 3.5.07**

D E C I S I O N
of Technical Board of Appeal 3.5.07
of 6 July 2022

Appellant: Google LLC
(Applicant)   1600 Amphitheatre Parkway
                            Mountain View, CA 94043 (US)

Representative: Williams, Michael David
                            Marks & Clerk LLP
                            1 New York Street
                            Manchester M1 4HD (GB)

Decision under appeal: **Decision of the Examining Division of the**
                            European Patent Office posted on 25 October 2018
                            refusing European patent application No.
                            14725807.3 pursuant to Article 97(2) EPC.

Composition of the Board:

Chairman   J. Geschwind
Members: C. Barel-Faucheux
                   M. Jaedicke

Summary of Facts and Submissions

I.   The appeal lies from the decision of the examining
         division to refuse European patent application No.
         14725807.3, which was filed as international
         application PCT/US2014/026226 (published as
         WO 2014/160282).

II.   The documents cited in the contested decision included:
         D1: Renqiang Min et al., "A Deep Non-Linear Feature
             Mapping for Large-Margin kNN Classification", 9th
             International Conference on Data Mining (ICDM),
             pp. 357-366, December 2009
         D2: WO 2009/111212 A2, published on 11 September 2009
         D3: Miklós Erdélyi et al., "Web Spam Classification: a
             Few Features Worth More", WebQuality'11:
             Proceedings of the 2011 Joint WICOW/AIRWeb
             Workshop on Web Quality, Hyderabad (India),
             pp. 27-34, 28 March 2011

III.   The examining division refused the application for lack
         of novelty of the subject-matter of claim 1 of the sole
         request over the disclosure of document D1 and lack of
         inventive step of the subject-matter of claim 1 having
         regard to D1, "if necessary" in combination with the
         disclosure of document(s) D2 and/or D3.

         Furthermore, according to the examining division, the
         features of claim 1 going beyond a well-known general
         purpose computer did not contribute to the technical
         character of the invention and therefore could not
         support the presence of an inventive step (reference
         was made to the Guidelines for Examination, G-VII,
         5.4).

The non-technical features did not contribute to the technical character of the invention because the purpose they served in the context of claim 1 was the automatic classification of resources into categories, which was clearly a non-technical task (reference was made to T 1784/06, Reasons 3.1; T 1358/09, Reasons 5.2; and T 2230/10, Reasons 3.4 and 3.11).

No technical use of the classification result was implied in claim 1. However, even if it were employed for recommending data items to a human user in response to a query as envisaged in the description of the application (page 4, last paragraph), this would not be a technical purpose either (reference was made to T 306/10, Reasons 5.2 and T 598/14, Reasons 2.4).

The examining division interpreted the appellant's announcement that it would not attend the oral proceedings as a withdrawal of the request for oral proceedings and exercised its discretion to take the decision without holding oral proceedings (following the Guidelines for Examination, E-III, 7.2.2).

IV.   With the statement of grounds of appeal, the appellant requested that the decision under appeal be set aside and that a patent be granted on the basis of a main request (corresponding to the set of claims refused by the examining division) or one of the first and second auxiliary requests filed with the statement of grounds.

V.   In a communication under Article 15(1) RPBA 2020 accompanying the summons to oral proceedings, the board expressed, among other things, its provisional opinion that the subject-matter of claim 1 of the main request and the first and second auxiliary requests did not appear to involve an inventive step (Article 56 EPC).

VI.   In reply to the communication, the representative
       stated that it would not be attending the oral
       proceedings and requested a decision "based on the
       file". Oral proceedings were then cancelled.

VII.   Claim 1 of the main request reads as follows:
       " A system (200) comprising:
           a deep network (206) implemented in one or more
       computers that defines a plurality of layers of non-
       linear operations, wherein the deep network comprises:
               an embedding function layer (208) configured to:
                   receive an input comprising a plurality of
       features (220) of a resource, wherein each feature is a
       value of a respective attribute of the resource, and
                   process each of the features (220) using a
       respective embedding function to generate one or more
       numeric values (222), and
               one or more neural network layers (210)
       configured to:
                   receive the numeric values (222), and
                   process the numeric values to generate an
       alternative representation (224) of the features of the
       resource, wherein processing the numeric values
       comprises applying one or more non-linear
       transformations to the numeric values; and
               a classifier (212) configured to:
                   process the alternative representation (224)
       of the input to generate a respective category score
       (226) for each category in a pre-determined set of
       categories, wherein each of the respective category
       scores measure a predicted likelihood that the resource
       belongs to the corresponding category."

VIII. Claim 1 of the first auxiliary request differs from
       claim 1 of the main request in that "for detecting

spam" has been inserted after "A system (200)" and that
the following text has been added at the end of the
claim:

"and wherein the pre-determined set of categories
includes a search engine spam category, and the
category score for the resource measures a predicted
likelihood that the resource is a search engine spam
resource".

IX.   Claim 1 of the second auxiliary request differs from
        claim 1 of the first auxiliary request in that
        "wherein each of the embedding functions is specific to
        features of a respective feature type, and wherein each
        of the embedding functions receives a feature of the
        respective type and applies a transformation to the
        feature that maps the feature into a numeric
        representation in accordance with a set of embedding
        function parameters,"
        has been inserted after the text
        "process each of the features (220) using a respective
        embedding function to generate one or more numeric
        values (222)".

X.   The appellant's arguments relevant to this decision are
        addressed in detail below.

Reasons for the Decision

The application

1.   The application relates to classifying search engine
        resources as a spam resource (or as belonging to the
        "spam" category) or not a spam resource (or as
        belonging to the "not spam" category).

2.   Figure 2 reproduced below illustrates this resource
        classification system 200:
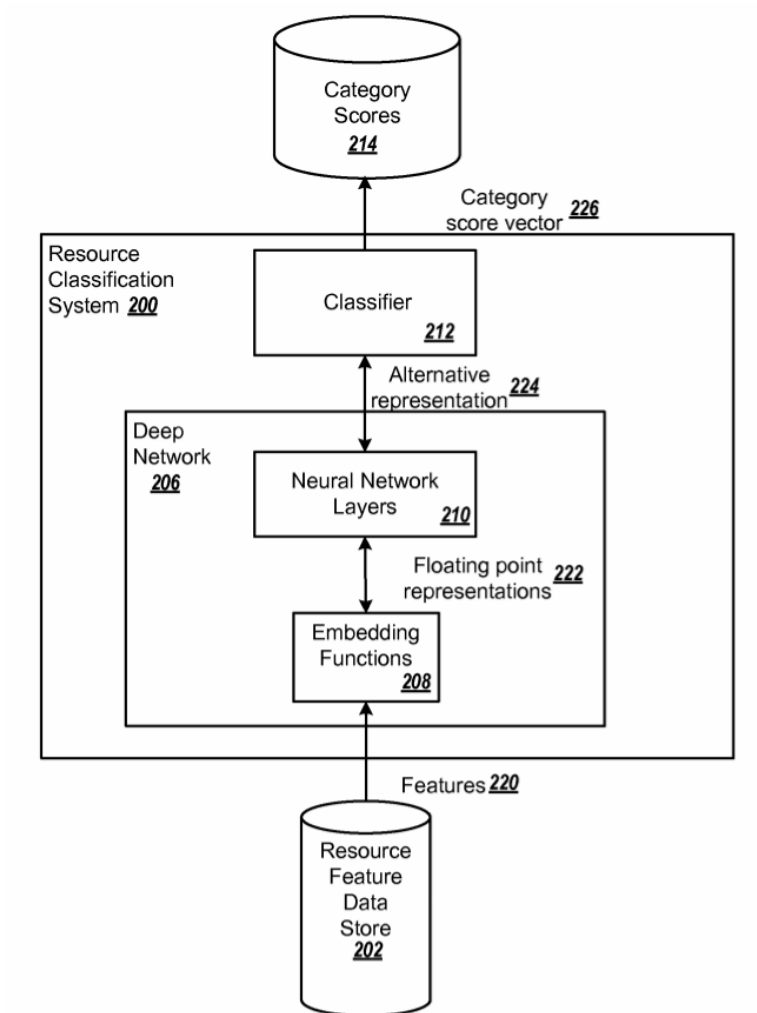
Main request - inventive step

3.   The board considers document D1 an appropriate starting
        point for assessing inventive step.

4.   In document D1, by combining the idea of deep learning
        and large-margin discriminative learning, a new kNN
        classification and supervised dimensionality reduction
        method called "DNet-kNN" is proposed. It encompasses a
        non-linear feature transformation to directly achieve

the goal of large-margin kNN classification, which is based on a deep encoder network with four hidden layers pre-trained with restricted Boltzmann machines (RBMs) (page 358, left-hand side, last paragraph; page 359, right-hand side, first paragraph; page 360, section "4 Large-margin kNN classification using deep neural networks", first paragraph and Figure 2).

5.   Document D1 discloses that "kNN" is one of the most popular classification methods due to its simplicity and reasonable effectiveness. It has been shown to have good performance for classifying many types of data (section "1 Introduction", page 357, left-hand column). The method of document D1 uses non-linear transformations so that each data point stays closer to its nearest neighbours having the same class as it than to any other data in the non-linearly transformed feature space (page 358, left-hand side, first paragraph).

6.   The method of document D1 is applied to classify newsgroup text data with binary features. This newsgroup text data constitutes a data set that contains binary occurrences for 100 words. The binary feature vectors have been used to classify the postings into four categories, which are "computer", "recreation", "science" and "talks" (section 5.3 on page 365). The board agrees with the examining division that an "embedding function" according to claim 1 is implicitly disclosed by this passage of document D1 since the newsgroup text data has to be transformed into numeric values at some stage.

7.   Thus, document D1 discloses a system comprising, in the language of claim 1 of the main request:

a deep network implemented in one or more
computers that defines a plurality of layers of non-
linear operations, wherein the deep network comprises:
an embedding function layer configured to:
receive an input comprising a plurality of
features of a resource, wherein each feature is a value
of a respective attribute of the resource, and
process each of the features using a respective
embedding function to generate one or more numeric
values, and
one or more neural network layers configured to:
receive the numeric values, and
process the numeric values to generate an
alternative representation of the features of the
resource, wherein processing the numeric values
comprises applying one or more non-linear
transformations to the numeric values; and
a classifier configured to process the alternative
representation of the input to generate *a plurality of
categories* (the part in italics being different from
the definition of the classifier in claim 1)

8.   The following are thus the distinguishing features of
claim 1 having regard to document D1.

The classifier according to claim 1 is configured to
process the alternative representation of the input to
generate <u>a respective category score (226) for each
category in a pre-determined set of categories, wherein
each of the respective category scores measure (*sic*) a
predicted likelihood that the resource belongs to the
corresponding category</u> (the underlined part being the
distinguishing features).

9.   The board notes that the application describes that the
categories might be:

- a "spam" category and a "not spam" category
  (description of the application as originally
  filed, page 5, lines 11 to 14)
- a "not spam" category and a category for each type
  of spam, such as "content spam", "link spam" and
  "cloaking spam", etc. (page 5, lines 29 to 34)
- a category for each resource type of a group of
  resource types, including "news resources", "blog
  resources", "forum resources", "shopping
  resources", "product resources" and "political
  resources" (page 6, lines 3 to 12)

The system generates a score for each category. A score
is a prediction of how likely it is that the resource
belongs to the corresponding category (page 5, lines 6
to 10; page 5, line 34 to page 6, line 2 and page 11,
line 31, to page 12, line 2). For example, the
generated score can be a predicted likelihood that the
resource is a "spam" resource (page 5, lines 26 to 28).

10.   However, it is apparent neither what a further
       technical effect of the distinguishing features could
       be, nor what objective technical problem the subject-
       matter of claim 1 would solve.

       In its statement of grounds of appeal, the appellant
       argued that following decision T 1227/05, the claimed
       system could at least be regarded as simulating a
       hardware circuit that classifies inputs and thus had a
       technical purpose. However, the board considers that
       decision T 1227/05 cannot be followed as argued by the
       appellant in view of recent decision G 1/19, Reasons
       133.

       In the absence of any technical effect beyond its
       straightforward implementation in one or more

computers, the subject-matter of claim 1 of the main request does not involve an inventive step (Article 56 EPC).

First auxiliary request – inventive step

11. Claim 1 of the first auxiliary request differs from claim 1 of the main request in that the system is specified as being a system "for detecting spam" and that the following text has been added at the end of the claim:

"and wherein the pre-determined set of categories includes a search engine spam category, and the category score for the resource measures a predicted likelihood that the resource is a search engine spam resource".

12. A "spam" category is one example of a category considered by the board as falling within the scope of claim 1 of the main request.

13. Since the board doubts that the meaning of the expression "search engine spam resource" is clear, it interprets this expression in light of the description for the assessment of inventive step. The application designates a search engine spam resource as a resource provided to a search system that has been manipulated by a spammer to give the resource a high search engine ranking as a response to one or more queries which the resource would not legitimately have. For example, content in a resource may be made <u>to appear particularly relevant to a specific geographic area</u>, and so be highly ranked for queries directed to that area, when in fact the content refers to a business, for example, that has no place of business in the area.

Search engine spam can include other forms of erroneous information as well (page 5, lines 15 to 23).

14.   The set of features added by the first auxiliary request relates to the categorisation of the resource as "spam" and appears to be merely of a non-technical cognitive nature. A "spam" resource could at most comprise intrinsic features related to the inherent result of its manipulation by a spammer. In any case, the board doubts that such a manipulation is technical. Even if it were, the result of the manipulation, i.e. the "spam(med) resource", would not necessarily be. Furthermore, claim 1 does not define how the alternative representation is processed to generate a predicted likelihood that the resource is a search engine spam resource. Generating a likelihood appears instead to be a mathematical feature, and not technical either.

15.   Therefore, the subject-matter of claim 1 of the first auxiliary request does not involve an inventive step (Article 56 EPC).

Second auxiliary request - inventive step

16.   Claim 1 of the second auxiliary request differs from claim 1 of the first auxiliary request in that "wherein each of the embedding functions is specific to features of a respective feature type, and wherein each of the embedding functions receives a feature of the respective type and applies a transformation to the feature that maps the feature into a numeric representation in accordance with a set of embedding function parameters," is inserted after

"process each of the features (220) using a respective
embedding function to generate one or more numeric
values (222)".

17.   These additional features in this added text are not
        disclosed by document D1.

18.   The effect of these additional distinguishing features
        (having regard to document D1) is to generate a numeric
        representation (i.e. one or more numeric values)
        according to the respective type of the resource
        features.

19.   In a similar case (T 872/19), the appellant argued, for
        the first auxiliary request comprising the same
        additional features, that the objective problem to be
        solved was to provide a more accurate representation of
        the resource features by generating an alternative
        representation of these resource features "in a new
        way".

        In decision T 872/19, Reasons 22.2, the board noted
        that providing a more accurate "alternative
        representation" of the resource features, and therefore
        refined relevance scores (the "category scores" in the
        current case), was not a technical effect since the
        relevance scores ("category scores" here) did not
        constitute a technical feature. The board considers
        that the same reasoning also applies to claim 1 of the
        second auxiliary request in the case at hand.

20.   In the absence of any technical effect beyond its mere
        implementation in one or more computers, the subject-
        matter of claim 1 of the second auxiliary request
        cannot be considered to involve an inventive step
        (Article 56 EPC).

Conclusion

21.   Since none of the requests is allowable, the appeal is
        to be dismissed.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar: The Chairman:

G. Nachtigall   J. Geschwind

Decision electronically authenticated