Trials@uspto.gov
571-272-7822

Paper 38
Entered: September 12, 2022

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

INTEL CORPORATION,
Petitioner,

v.

HEALTH DISCOVERY CORPORATION,
Patent Owner.

IPR2021-00552
Patent 7,542,959 B2

Before LYNNE H. BROWNE, GARTH D. BAER, and FREDERICK C. LANEY, *Administrative Patent Judges*.

Opinion for the Board filed PER CURIAM.

Opinion Dissenting filed by Administrative Patent Judge BAER.

PER CURIAM.

JUDGMENT
Final Written Decision
Determining No Challenged Claims Unpatentable
35 U.S.C. § 318(a)

I. INTRODUCTION

Intel Corp. ("Petitioner") filed a Petition to institute an *inter partes* review of claims 1, 5, 6, 10, 12, and 16 (the "challenged claims") of U.S. Patent 7,542,959 B2 (Ex. 1001, the "'959 patent"). Paper 3 ("Petition" or "Pet."). Health Discovery Corp. ("Patent Owner") filed a Preliminary Response. Paper 11 ("Prelim. Resp."). On September 14, 2021, we instituted trial. Paper 21 ("Inst. Dec."). Patent Owner filed a Response to the Petition (Paper 24, "PO Resp."), Petitioner filed a Reply (Paper 27, "Reply"), and Patent Owner filed a Sur-reply (Paper 29, "Sur-reply"). We held an oral hearing on June 14, 2022, and the hearing transcript is included in the record. *See* Paper 37 ("Tr.").

We have jurisdiction to conduct this *inter partes* review under 35 U.S.C. § 6. This Final Written Decision is issued pursuant to 35 U.S.C. § 318(a) and 37 C.F.R. § 42.73. For the reasons below, we determine that Petitioner has not shown by a preponderance of evidence that claims 1, 5, 6, 10, 12, and 16 of the '959 patent are unpatentable.

A. RELATED MATTERS

The parties identify the following related proceeding involving the '959 patent: *Health Discovery Corp. v. Intel Corp.*, Civil Action No. 6:20-cv-666 (W.D. Texas July 23, 2020). Pet. 74; Paper 4 (Patent Owner's Mandatory Notices) 2.

B. THE '959 PATENT

The '959 patent addresses using Support Vector Machines (SVM) and Recursive Feature Elimination (RFE) for selecting genes that are capable of accurately distinguishing between medical conditions. Ex. 1001, code (57). Generally, the '959 patent describes the identification of a determinative subset of features from within a large set of features is performed by training an SVM to rank the features according to classifier weights, where features are removed to determine how their removal affects the value of the classifier weights. *Id.* "The features having the smallest weight values are removed and a new support vector machine is trained with the remaining weights." *Id.* "The process is repeated until a relatively small subset of features remain that is capable of accurately separating the data into different patterns or classes." *Id.* Figure 2 is reproduced below.

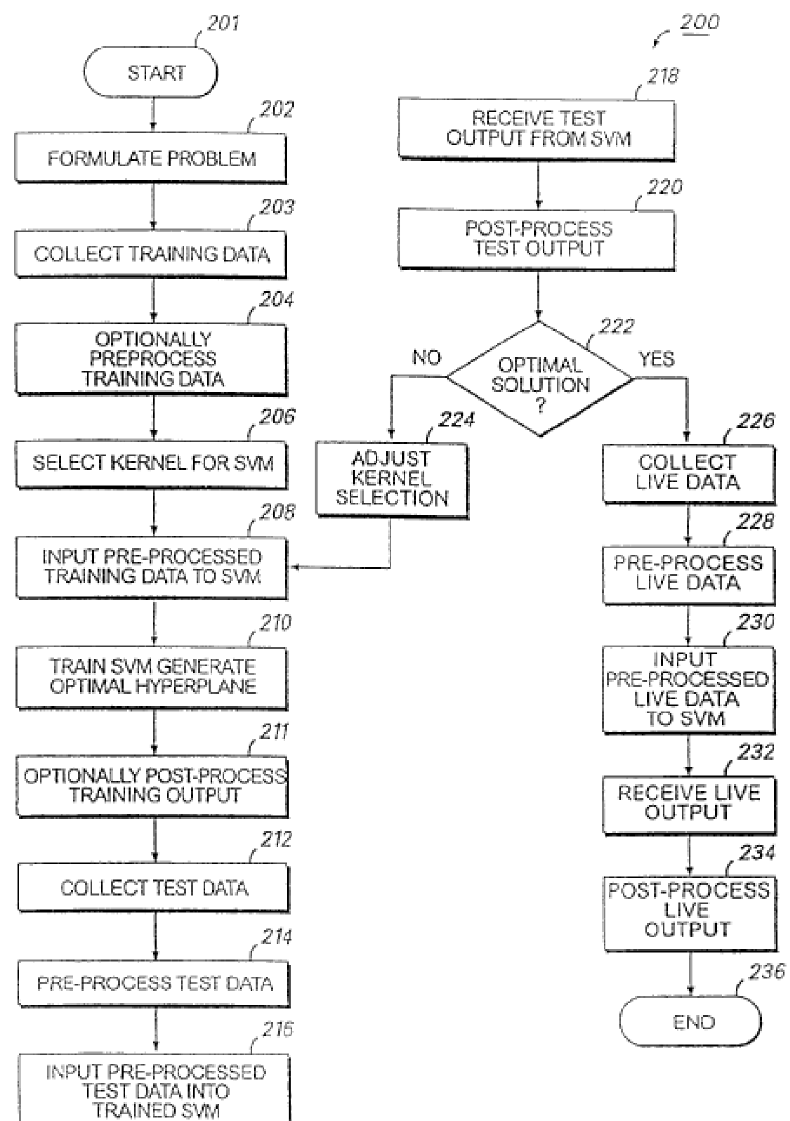

FIG.2

Figure 2 is a flowchart illustrating an example method for increasing knowledge that may be discovered using a learning machine. *Id.* at 7:63–65. As shown in Figure 2, the SVM is trained using training data to generate an optimal hyperplane. *Id.* at 16:51–17:4. Test data is input into the trained SVM "to determine whether the SVM was trained in a desirable manner." *Id.* at 17:11–13. If not, the kernel selection is adjusted at step 224 and the training process is repeated from step 208. *Id.* at 16:47–57. After the optimal kernel is selected, the SVM is further optimized through feature selection to reduce the dimensionality of feature space. *See id.* at 26:20–33. The '959 patent uses RFE, where the feature corresponding to the smallest weight in the new classifier is eliminated and, at each iteration, a new classifier is trained with the remaining features. *Id.* at 52:52–64.

C. CHALLENGED CLAIMS

Petitioner challenges claims 1, 5, 6, 10, 12, and 16 of the '959 patent. Pet. 1. Challenged claims 1, 12, and 16 are independent. Claim 1 is reproduced below.

> 1. A computer-implemented method for predicting patterns in biological data, wherein the data comprises a large set of features that describe the data and a sample set from which the biological data is obtained is much smaller than the large set of features, the method comprising:
>
> identifying a determinative subset of features that are most correlated to the patterns comprising:
>
> (a) inputting the data into a computer processor programmed for executing support vector machine classifiers;
>
> (b) training a support vector machine classifier with a training data set comprising at least a portion of the sample set and having known outcomes with respect to the patterns, wherein the classifier

comprises weights having weight values that correspond to the features in the data set and removal of a subset of features affects the weight values;

(c) ranking the features according to their corresponding weight values;

(d) removing one or more features corresponding to the smallest weight values;

(e) training a new classifier with the remaining features;

(f) repeating steps (c) through (e) for a plurality of iterations until a final subset having a pre-determined number of features remains; and

generating at a printer or display device a report comprising a listing of the features in the final subset, wherein the final subset comprises the determinative subset of features for determining biological characteristics of the sample set.

Ex. 1001, 62:6–35.

D. ASSERTED GROUNDS OF UNPATENTABILITY

Petitioner asserts the following grounds of unpatentability. Pet. 3. Petitioner submits the Declarations of Dr. Theodoros Evgeniou (Exs. 1003, 1095) in support of its arguments.

Claims Challenged	35 U.S.C. §[1]	References
1, 5, 6, 10, 12, 16	103	Kohavi[2], Boser[3], Hocking[4]

[1] The Leahy-Smith America Invents Act ("AIA") amended 35 U.S.C. § 103. *See* Pub. L. No. 112-29, 125 Stat. 284, 285–88 (2011). As the application that issued as the '959 patent was filed before the effective date of the relevant amendments, the pre-AIA version of § 103 applies.

[2] Kohavi et al., "Wrappers for Feature Subset Selection," Artificial Intelligence 97, 273-324 (1997) (Ex. 1007).

[3] US Patent No. 5,649,068, July 15, 1997 (Ex. 1008).

Claims Challenged	**35 U.S.C. §[1]**	**References**
1, 5, 6, 10, 12, 16[5]	103	Kohavi, Boser, Hocking, Cristianini[6]

Patent Owner counters with the Declaration of Dr. Clayton Scott (Ex. 2014).

II. DISCUSSION

A. LEGAL STANDARDS

Petitioner has the burden of proof for establishing the challenged claims are unpatentable. *Harmonic Inc. v. Avid Tech., Inc.*, 815 F.3d 1356, 1363 (Fed. Cir. 2016) ("In an IPR, the petitioner has the burden from the onset to show with particularity why the patent it challenges is unpatentable.").

Section 103(a) forbids issuance of a patent when "the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains." *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of underlying factual determinations, including: (1) the scope and content of the prior art; (2) any differences between the claimed subject matter and the prior art; (3) the level of ordinary skill in the art; and (4) when available, evidence such as commercial success, long felt but unsolved needs, and failure of

---

[4] Hocking et al., "Selection of the Best Subset in Regression Analysis," Technometrics, 9:4, 531–540 (1967) (Ex. 1009).

[5] Although in the table provided on page 3 of the Petition only identifies claim 10 for Ground 2, the Petition at pages 64–69 addresses claims 1, 5, 6, 10, 12, and 16.

[6] Cristianini, N., et al., "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods," Cambridge University Press (2000) (Ex. 1010).

others. *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966); *see KSR*, 550 U.S. at 407 ("While the sequence of these questions might be reordered in any particular case, the [*Graham*] factors continue to define the inquiry that controls."). The Court in *Graham* explained that these factual inquiries promote "uniformity and definiteness," for "[w]hat is obvious is not a question upon which there is likely to be uniformity of thought in every given factual context." 383 U.S. at 18.

The Supreme Court made clear that we apply "an expansive and flexible approach" to the question of obviousness. *KSR*, 550 U.S. at 415. Whether a patent claiming the combination of prior art elements would have been obvious is determined by whether the improvement is more than the predictable use of prior art elements according to their established functions. *Id.* at 417. To support this conclusion, however, it is not enough to show merely that the prior art includes separate references covering each separate limitation in a challenged claim. *Unigene Labs., Inc. v. Apotex, Inc.*, 655 F.3d 1352, 1360 (Fed. Cir. 2011). Rather, obviousness additionally requires that a person of ordinary skill at the time of the invention "would have selected and combined those prior art elements in the normal course of research and development to yield the claimed invention." *Id.*; *see also Orexo AB v. Actavis Elizabeth LLC*, 903 F.3d 1265, 1273 (Fed. Cir. 2018) ("The question is not whether the various references separately taught components of the '330 Patent formulation, but whether the prior art suggested the selection and combination achieved by the '330 inventors.").

In determining whether there would have been a motivation to combine prior art references to arrive at the claimed invention, it is insufficient to simply conclude the combination would have been obvious without identifying any reason *why* a person of skill in the art would have made the combination. *Metalcraft of Mayville, Inc. v. Toro Co.*, 848 F.3d 1358, 1366 (Fed. Cir. 2017). As factfinders, we also must be aware "of the distortion caused by hindsight bias and must be cautious of arguments reliant upon *ex post* reasoning." *KSR*, 550 U.S. at 421.

Applying these general principles, we consider the evidence and arguments of the parties.

B. LEVEL OF SKILL IN THE ART

Petitioner contends a person of ordinary skill in the art at the time of the alleged invention of the '959 patent would have had "at least a Master's degree in electrical engineering, computer science, or the equivalent with three years of experience in machine learning and data-analysis techniques." Pet. 21 (citing Ex. 1003 ¶ 15). Further, "[a]dditional education could substitute for professional experience, and vice versa. *Id*. Patent Owner does not contest that a skilled artisan would have this kind background knowledge and experience, except that it "believes that the use of the term 'at least' is inappropriate inasmuch as it would extend to persons having more than ordinary skill in the art." PO Resp. 22–23. We agree with Patent Owner and adopt the parties' articulation of the level of ordinary skill in the art without the term "at least." This articulation is supported by Dr. Evgeniou's testimony and appears commensurate with the level of ordinary skill as reflected in the asserted prior art and the '959 patent.

C. CLAIM CONSTRUCTION

Petitioner proposes that we construe the term "data" as "biological data." Pet. 20–21. Patent Owner disputes Petitioner's construction, and asserts that we need not construe the term to resolve any asserted ground in this petition. *See* PO Resp. 19–22. We agree with Patent Owner that no claim terms require express construction in order to determine whether or not to institute *inter partes* review. *See Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) ("[W]e need only construe terms 'that are in controversy, and only to the extent necessary to resolve the controversy.'") (quoting *Vivid Techs., Inc. v. Am. Sci. & Eng'g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999)).

D. ASSERTED PRIOR ART

1. Kohavi (Ex. 1007)

Kohavi is titled "Wrappers for Feature Subset Selection." Kohavi teaches a feature subset selection method for selecting a relevant subset of features upon which to focus a learning algorithm's attention, while ignoring the rest. Ex. 1007, abstract.

2. Boser (Ex. 1008)

Boser teaches a "pattern recognition system using support vectors"—i.e., an SVM. Ex. 1008, code (54). Boser's Figure 5 is reproduced below.

FIG. 5

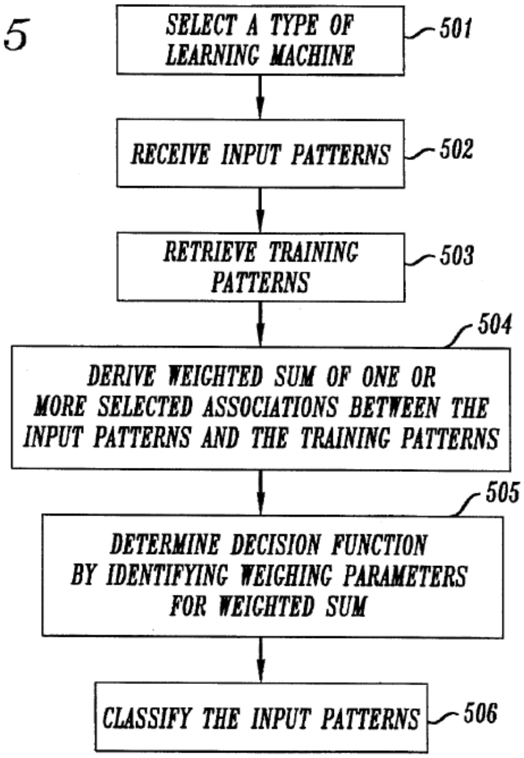

Figure 5 is a flow diagram that implements the principles of Boser's SVM. *Id.* at 3:7–9.

3. Hocking (Ex. 1009)

Hocking is titled "Selection of the Best Subset in Regression Analysis." Ex. 1009, 531. Hocking teaches an iterative process that removes variables based on weight-vector ranking, until a subset that provides the best regression is identified. Ex. 1003 ¶¶ 204–210 (citing Ex. 1009, 531, 533).

E. GROUND 1 UNPATENTABILITY ANALYSIS

In its first ground of unpatentability, Petitioner contends that claims 1, 5, 6, 10, 12, and 16 would have been obvious in view of Kohavi, Boser, and Hocking. Pet. 21–64. Patent Owner alleges that several dispositive deficiencies exist with Petitioner's contention; in particular, Patent Owner argues that Petitioner has failed to establish the recited computer-implemented method for identifying patterns in data including "ranking the features according to their corresponding weight values" (PO Resp. 26–51), as each independent claim challenged recites (Element 1(c), Element 12(c), Element 16(c) (Ex. 1001, 62:23–24, 63:15–16, 64:18–19)). Because we agree that this issue is dispositive and because we find Petitioner's evidence unpersuasive for the reasons explained below, we find Petitioner has not made an adequate showing that independent claims 1, 12, and 16 are unpatentable. Therefore, our discussion below focuses on Elements 1(c), 12(c), and 16(c).

1. Petitioner Contends Kohavi, Boser, and Hocking Disclose Ranking the Features According to Their Corresponding Weight Values

For Element 1(c), Petitioner contends that Hocking discloses the use of estimated weight values as ranking criteria in a regression analysis. Pet. 46–48 (citing Ex 1009, 531–533; Ex. 1003 ¶¶ 246, 249). And in view of the combined teachings of Kohavi, Boser, and Hocking, Petitioner contends that "at a given iteration of Kohavi's wrapper method, a feature subset is evaluated using Boser's SVM as the induction algorithm, and ranking criteria [are] computed based on the estimates of feature weight values determined according to the teachings of Hocking." *Id.* at 48 (citing Ex. 1003 ¶ 251). Petitioner concludes that the combination discloses "ranking the features according to their corresponding weight values." *Id.* (citing Ex. 1003 ¶¶ 246–251).

Petitioner contends that a skilled artisan would have known to select the SVM that Boser discloses as the induction algorithm in the wrapper method that Kohavi discloses. Pet. 27. Petitioner contends that because Kohavi teaches a wrapper method that is used with an induction algorithm, "which is a type of machine learning that induces a general rule from a set of observed instances," a skilled artisan would have been led to Boser's SMV. *Id.* (citing Ex. 1007 ¶ 274; Ex. 1003 ¶ 211). This is so, according to Petitioner, because Boser is directed to a "learning machine" and a skilled artisan would have viewed Boser's SVM as a kind of induction algorithm that Kohavi references because "it formulates a decision function from a set of training examples." *Id.* at 28 (citing Ex. 1008, 1:9, 1:12–19; Ex. 1003 ¶ 211).

Petitioner contends that Kohavi teaches that the wrapper method may may "achieve better generalization with reduced computational burden," which was a known problem at the time of the invention with "practical learning scenarios." Pet. 28 (citing Ex. 1003 ¶¶ 212, 276). Petitioner notes, for example, Kohavi states that "[s]ignificant improvement in accuracy is achieved for some datasets for the two families of induction algorithms used: decision trees and Naïve-Bayes." *Id.* (citing Ex. 1007, 273 (Abstract)). And a skilled artisan would have "recognized that, like SVMs, decision trees and Naïve-Bayes algorithms are types of supervised machine learning" and, therefore, Petitioner contends Kohavi suggests to a skilled artisan that similar improvements reducing computational burden may be gained by using Boser's SVM with Kohavi's wrapper method. *Id.* at 28–29 (citing Ex. 1003 ¶¶ 212).

Combining Boser and Kohavi in this manner would yield predictable results, according to Petitioner, because "[b]oth are specifically directed to achieving optimal classification results." Pet. 29 (citing Ex. 1008, 1:62–2:4; Ex. 1003 ¶¶ 213, 274). Additionally, Petitioner contends that a skilled artisan would have had a reasonable expectation of success because Kohavi teaches that the benefits of the disclosed wrapper method "are applicable to any type of induction algorithm" and, as a result, Kohavi's wrapper method would have been known to improve Boser's SVM "in the same ways as taught by Kohavi." *Id.* at 30 (citing Ex. 1003 ¶ 214; Ex. 1007, 274; Ex. 1008, 4:4–5).

Petitioner contends further that a skilled artisan would have known to further modify the Kohavi/Boser combination with "Hocking's simple and computationally efficient ranking based on weight values." Pet. 31. In particular, Petitioner contends that a skilled artisan "would have been motivated to modify the feature selection evaluation to use vector weights of the training model as ranking criteria, as taught by Hocking." *Id.* at 30–31 (citing Ex. 1003 ¶ 215). This modification, according to Petitioner, would require "[n]o substantive modification" to the Kohavi/Boser combination "other than the user of Boser's feature weight values as the ranking criteria." *Id.* at 31 (citing Ex. 1003 ¶ 215). Petitioner states that, "[w]hile Kohavi uses accuracy as ranking criteria, Kohavi explains that 'one can trivially use a cost function instead of accuracy as the evaluation function for the wrapper'" and "using Boser's feature weight values as the evaluation function of Kohavi's wrapper method, according to the teachings of Hocking, would be a trivial substitution of known methods." *Id.* (citing Ex. 1003 ¶¶ 215, 309; Ex. 1007, 309).

"Hocking motivates the use of its ranking criteria" with the Kohavi/Boser combination, Petitioner contends, because Kohavi's wrapper method may use "backward feature elimination, which was known to be computationally complex" and Hocking teaches that "us[ing] the weight vectors of its linear model to identify the subset of variables that provides the 'best regression'" can be done "with **a minimum of computation**." Pet. 31 (citing Ex. 1007, 276, 284–285; Ex. 1009, 531, 533; Ex. 1003 ¶ 216). Petitioner states additionally that "Hocking's ranking criteria using vector weights lends itself to use in Boser's SVM, because Boser's derived decision function maximizes margin by optimizing weight vectors associated with input patterns." *Id.* at 32 (citing Ex. 1008, 5:17–25; Ex. 1003, ¶ 216).

Petitioner contends that combining Hocking with Kohavi and Boser "merely combines prior art elements according to known methods to yield predictable results and applies a known technique (Hocking's variable selection) to a known device (Kohavi's [Wrapper] method using Boser's SVM) which is ready for improvement to yield predictable results." Pet. 32 (citing *KSR*, 550 U.S. at 417; Ex. 1003 ¶ 217). To support this contention, Petitioner highlights that Hocking and Boser's SVM both utilize "classical statistics principles" and are both directed to types of linear regressions that use similar linear modeling. *Id.* (citing Ex. 1008, 5:17–22; Ex. 1009, 531; Ex. 1003 ¶ 217). Therefore, Petitioner concludes, combining the teachings of Hocking with the SVM that Boser teaches would yield predictable results. *Id.*

In addition, Petitioner contends that a skilled artisan would have had a reasonable expectation of success combining Kohavi, Boser, and Hocking. Pet. 32 (citing Ex. 1003 ¶ 218). According to Petitioner, "Kohavi's recursive, backwards elimination techniques are based on the same classical statistics principles described in Hocking's iterative variable evaluation method" and "the proposed modification to Kohavi's [wrapper method], uses the known technique of Hocking's classical statistical method to improve Kohavi's similar [wrapper] method in the same way." *Id.* at 32–33 (citing Ex. 1003, ¶ 218; Ex. 1007, 314–315; Ex. 1028, 505).

2. Patent Owner's Response

Patent Owner asserts that Petitioner's challenge fails because there is an insufficient basis to support a finding that, at the time of the invention, a skilled artisan would have known to modify the feature selection search module in Kohavi's wrapper method to compute ranking criteria using the optimized plurality of weights generated by Boser's SVM. *See* PO Resp. 26–51.

Patent Owner first argues that "Petitioner's proposed modification of Kohavi, or Kohavi-Boser, based on the teachings of Hocking represent . . . a marked change in the principle of operation of Kohavi (or Kohavi-Boser) inasmuch as it vitiates a fundamental precept of the wrapper approach as taught by Kohavi." PO Resp. 26. Patent Owner argues that a feature selection search module capable of performing its tasks using an induction algorithm, but without any knowledge of the algorithm itself, is a principle of operation for the wrapper method Kohavi discloses. *See id.* 26–37. Patent Owner asserts that "[t]he wrapper approach is so-named because it wraps a search for and evaluation of possible feature sets around an induction algorithm, which itself is used as part of the evaluation operation." *Id.* at 27 (citing Ex. 1007, 284). Patent Owner notes that, in Kohavi's words, "[t]he idea behind the wrapper approach . . . is simple: ***the induction algorithm is considered as a black box***." *Id.* at 26 (citing Ex. 1007, 274). Moreover, Patent Owner notes that, when explaining the wrapper method, Kohavi begins by reiterating that, "[i]n the wrapper approach, . . . the feature subset selection is done using the induction algorithm as a black box (i.e., no knowledge of the algorithm is needed, just the interface)." *Id.* at 27 (citing Ex. 1007, 284).

Patent Owner asserts that "[b]ecause the wrapper approach allows the induction algorithm to be treated as a black box, the feature selection module need not be concerned with the particulars of the induction algorithm" and the feature selection module is able to perform its tasks (i.e., searching for candidate feature sets and selecting a feature set with the highest evaluation) without any direct interaction with the induction algorithm. PO Resp. 30–31 (citing Ex. 1007, 284); *see also id.* 28–30 (explaining that Kohavi's feature selection search module performs two primary tasks: (1) "it *searches* for candidate feature sets to evaluate" and (2) it "*selects* one of the candidate feature sets with the highest evaluation as the final set on which to run the induction algorithm") (citing Ex. 1007, 274, 284, Figs 1, 5; Ex. 2013, 34:2–10)). "Instead, all that the two . . . primary tasks of the feature selection search module needs to carry out its primary tasks is the *interface* to the induction algorithm (i.e., the "feature evaluation" module [illustrate in Figure 1 of Kohavi])." *Id.* at 31 (citing Ex. 1007, 284, 274 (Fig. 1)).

Patent Owner argues that "in the [P]etition's application of Hocking to the combination of Kohavi and Boser, the [P]etition proposes to *expose* the contents of the black box to the feature selection search module, changing the principle of operation of Kohavi." PO Resp. 32. Patent Owner argues that "rather than following Kohavi's direction and treating the induction algorithm (Boser's SVM in the proposed combination) as a 'black box,' Petitioner now demands that the induction algorithm be exposed to the feature selection search module, and in particular that the feature selection search module have knowledge of the weights of the SVM decision function." *Id.* Patent Owner argues that "[b]y doing so, the induction algorithm can no longer be considered, and it no longer is treated as, 'a black box.'" *Id.* at 35. As a result, Patent Owner argues "[a] modification of Kohavi (or of Kohavi-Boser) of this magnitude is not a mere combination of prior art elements;" instead "[i]t is nothing less than a change in the principle of operation of Kohavi (or Kohavi-Boser) and one of such a character that it undermines" Petitioner's proposed combination. *Id.* at 36.

Patent Owner argues next that Petitioner's rationale for modifying the Kohavi/Boser combination in view of the ranking technique Hocking discloses lacks support because Hocking (as well as Kohavi or Boser) does not suggest the weights created by Boser's SVM are interchangeable for the estimated weight values that Hocking discloses. *See* PO Resp. 39–44. Patent Owner asserts that Hocking teaches a ranking technique that depends on determining estimated feature weights based on distribution assumptions (i.e., "linear model assumptions") that are unsuitable for SVM applications and, at the time of the invention, similar suitable distribution assumptions for SVMs were not known to skilled artisans. *Id.* at 39–42 (citing Ex. 1009, 531, 533; Ex. 2014 ¶¶ 34–37; Ex. 2013, 70:16–71:21). Furthermore, Patent Owner asserts that a skilled artisan would have recognized that the estimated feature weights calculated in the manner taught by Hocking would generate different values than the ones generated through the optimization of Boser's SVM. *Id.* at 43–44 (citing Ex. 2014 ¶¶ 35–36). "[N]or is there any known mathematical formula that can be used to convert the former to the later," according to Patent Owner. *Id.*

Additionally, Patent Owner asserts Hocking and Boser teach eliminating features to serve different objectives. PO Resp. 44–47. Patent

Owner asserts that the objective of feature elimination in Hocking is to minimize the "residual sum of squares" and "the objective considered in Boser is based on geometric considerations that are entirely unconnected with the 'residual sum of squares.'" *Id.* at 44–45 (citing Ex. 2014 ¶ 35; Ex. 1009, 533; Ex. 2013, 52:14–53:1). According to Patent Owner, the objective of feature elimination in Boser is to optimize the margin used for classification in an SVM. *Id.* at 45–47 (citing Ex. 1008, 1:27–41; Ex. 2013, 64:3–20). Patent Owner argues that a skilled artisan "considering Boser and Hocking in their respective entireties would encounter, recognize, and appreciate the differing objectives of Boser and Hocking, and, consequently, would not have been motivated to apply the feature elimination teachings of Hocking to any Kohavi-Boser combination." *Id.* at 47.

Lastly, Patent Owner argues that Petitioner's contention that the known computational complexities associated with backwards feature elimination would have inspired a skilled artisan to combine Kohavi and Hocking is unsupported because Hocking does not teach backward elimination. PO Resp. 47–48. Instead, "Hocking teaches something different." *Id.* at 48. Patent Owner asserts that "Hocking's method does not begin in an initial state consisting of all features, and does not repeatedly (and permanently) remove one or more features in an iterative fashion." *Id.* Because "Hocking is focused on computing the globally optimal subset of features for the least squares objective," Patent Owner asserts that Hocking teaches that "for any value of p (a targeted feature subset size) it may be necessary that all of the . . . regressions be evaluated." *Id.* (citing Ex. 2014 ¶ 58; Ex. 1009, 532, 534). Patent Owner asserts that in Hocking, because "the feature subsets size r does not depend on the stage q," "when progressing to the next stage, Hocking does not increase 'the number of features to be removed' (r), but rather increases the number of features (= 'variables' in Hocking terminology) from which a feature subset of size r is sought." *Id.* (citing Ex. 2014 ¶ 59).

3. Petitioner's Reply

In Reply, Petitioner criticizes Patent Owner's argument regarding Kohavi's principle of operation as based on an erroneous characterization of the operation and Kohavi's teaching of the induction algorithm being a "black box." Reply 2–6. Petitioner contends that Kohavi's treatment of the induction algorithm as a "black box" is *not* an essential aspect of its operation. *Id.* at 2. Although "[t]he ability to use any induction algorithm in Kohavi . . . is a feature of Kohavi's wrapper method," Petitioner contends the "principle of operation is much more." *Id.* Petitioner instead points to Kohavi describing a "wrapper method [that] searches for an optimal feature subset tailored to a particular algorithm and a domain." *Id.* at 2–3 (citing Ex. 1007, 273 (Abstract)). Petitioner contends that "[w]hen deployed in the Kohavi-Boser-Hocking combination, the modified Kohavi functions as intended—it selects a subset of features, using the output of the induction algorithm and 'the induction algorithm itself as part of the function evaluating feature subsets,' all without having to know the underlying details of the induction algorithm." *Id.* at 3 (citing Ex. 1007, 274; Ex. 1003, ¶¶ 255, 268, 271; Pet. 23–25, 29–31). And, thus, Petitioner concludes "Kohavi operates on the same principles before and after the combination; the proposed combination does not change the principle of operation of Kohavi." *Id.*

Petitioner also criticizes Patent Owner's principle of operation argument as "twist[ing] the meaning" of the straightforward concept of a "black box" to reach the "tortured conclusion" that "all Kohavi's 'feature selection module needs to carry out its primary tasks is the interface to the induction algorithm (i.e., the 'feature evaluation' module).'" Reply 4. Petitioner contends, to the contrary, that "Kohavi makes clear that its '**feature subset selection algorithm** conducts a search for a good subset **using the induction algorithm itself** as part of the function evaluating feature subsets'" and, thus, "Kohavi's feature subset selection encompasses the feature selection search, feature evaluation, and induction algorithm depicted in Figure 1." *Id.* at 4–5 (citing Ex. 1095, ¶¶ 14–15; Ex. 1007, 274).

Petitioner contends further that, contrary to Patent Owner's suggestion that Kohavi's feature subset selection is "ignorant of the induction algorithm," Kohavi "is simply agnostic as to what induction algorithm a practitioner selects to use with it." Reply 5. It is in this manner, Petitioner contends, "[t]he induction algorithm is used as a 'black box' by the subset selection algorithm" of the wrapper method." *Id.* (citing Ex. 1007, 274). Petitioner contends that "Kohavi's wrapper method is interoperable with any selected induction algorithm and requires no knowledge of the inner workings of the algorithm, just the interface, so it can send inputs (feature set) and receive an output (a hypothesis) from the algorithm." *Id.* (citing Ex. 1007, 274 (Figure 1); Ex. 1003 ¶¶ 198–99; Pet. 23). Kohavi "explicitly illustrates this concept [by] showing that the induction algorithm outputs the algorithm's hypothesis as part of the evaluation process of the feature subset selection algorithm" and a skilled artisan would understand, Petitioner contends, that an induction algorithm's hypothesis includes information such as weights associated with the set of features input to the induction algorithm. *Id.* (citing Ex. 1095, ¶¶ 15–17; Ex. 1094, 55:4–8).

For Patent Owner's principle of operation arguments, Petitioner contends that "using Boser's feature weight values, which are part of the hypothesis output from the induction algorithm—a fact agreed upon by both experts—in the combination is consistent with Kohavi and does not change its principle of operation." Reply 7. It is consistent, Petitioner reasons, because Kohavi discloses that "one can trivially use a cost function instead of accuracy as the evaluation function for the wrapper," and "using Boser's feature weight values, estimated by optimizing a cost function, for evaluation with Hocking's ranking principle in Kohavi's feature subset selection would likewise be a trivial substitution of known methods." *Id.* (citing Ex. 1007, 309; Pet. 31; Ex. 1003 ¶ 215; Ex. 2013, 76:20–77:19).

Next in the Reply, Petitioner disputes Patent Owner's assertion that the assumptions and objectives underlying Hocking's linear regression cut against the proposed Kohavi/Boser/Hocking combination. Reply 9–16. Petitioner contends that Patent Owner's assertions are flawed because they rely on bodily incorporation, which is not a legal requirement for establishing obviousness. *Id.* at 9. Additionally, Petitioner attacks Patent Owner's assertions as incorrectly suggesting that "SVMs and linear regression are fundamentally disparate concepts such that a POSITA would not be motivated to apply Hocking's ranking principles to Boser's feature weights as part of Kohavi's wrapper method." *Id.* at 9–10 (citing Ex. 1003, ¶ 217). Petitioner notes that both experts have agree similarities exist between linear regression and SVMs and have acknowledged "that these techniques can be used to solve the same types of problems." *Id.* at 10 (citing Ex. 1094, 54:13–15, 45:18–46:2; Ex. 1003 ¶¶ 38, 66–68, 274). Petitioner also points out that Patent Owner's expert "has taught SVMs, linear regression, and Kohavi's wrapper methods to students in a single graduate course." *Id.* (citing Ex. 1094, 34:10–22, 48:9–19).

Even accepting that the same assumptions do not apply to the linear modeling for SVMs as are applied in Hocking, Petitioner contends that "it does not follow that a [skilled artisan] would have turned a blind eye to all references relating to linear regression (or to the field of classical statistics, more generally) when working in SVM-based machine learning." Reply 10. Instead, Petitioner contends that a skilled artisan "would have understood the differences between SVMs and linear regression and would have applied that knowledge when combining teachings from publications discussing linear regressions, such as Hocking, with an SVM-based algorithm, like the Kohavi-Boser combination." *Id.* at 10–11 (citing Ex. 1095 ¶¶ 24–25). Thus, Petitioner concludes, "when considering linear regression teachings for SVM-based machine learning, a [skilled artisan] would consider and apply reasonable assumptions with respect to those concepts . . . when combining concepts from those models." *Id.* at 11 (citing Ex. 1095, ¶ 24).

Petitioner states that Patent Owner's observations about the inaccuracies and lack of mathematical precision amounts to nothing more than "quibbl[ing]." Reply 11. Petitioner contends that a skilled artisan would have understood that the error and error variance concepts of linear regression do not exist in SVMs and, therefore, would have understood also that, for purposes of applying Hocking's teachings to an SVM, the associated variable could be "disregarded, resulting in the ranking criteria simplifying to a function of feature weights." *Id.* at 11–12 (citing Ex. 1095,

¶¶ 25–26; Ex. 1094, 56:19–57:21; Pet. 42; Ex. 1003 ¶ 249; Ex. 2013, 52:18–54:2). "And as explained in the Petition, a [skilled artisan] would understand that Hocking's regression weight estimates . . . are analogous to Boser's SVM weight vectors." *Id.* at 12 (citing Pet. 26–27, 46–48).

Petitioner contends that "the concepts and equations described in Hocking and in SVM publications, such as Boser, are strikingly similar" and, "[w]hile Hocking and Boser may not share the exact same 'linear model assumptions,' both disclose compatible linear models and are described using equations that are based on and include the equation of a line." Reply 12–13 (citing Pet. 24–27, 30–33, 46–47; Ex. 1001, 31:1–8; Ex. 1003 ¶¶ 37–39, 203, 205, 217, 247; Ex. 1008, 1:27–31; 3:12–16, 4:24–35, 5:9–10; Ex. 1009, 531). From this, Petitioner concludes that a skilled artisan "would have understood the applicability of the ideas in Hocking to Boser's machine learning algorithm, and due to the substantial similarities the POSITA would have had a reasonable expectation of success when applying Hocking's ranking principles to Boser's feature weights." *Id.* at 13–14 (Citing *KSR*, 550 U.S. at 417; Pet. 30–33; Ex. 1003 ¶¶ 215–218).

Petitioner contends the different objectives between Hocking and Boser highlighted by Patent Owner "should be disregarded because it compares two completely different concepts, which necessarily have two different objectives." Reply 15. Petitioner contends that, because Hocking performs "feature elimination" and Boser optimizes feature weights of its classification algorithm's decision function, "[t]he two objectives of Hocking and Boser are therefore complementary." *Id.* (citing Ex. 1003 ¶ 217). Petitioner states that Patent Owner "misunderstands the Petition" and asserts,

> the combination's "objective" and "optimization criteria" through which feature weights are computed are those of Boser, since SVM is the combination's classifier . . . Put simply, the Petition does not argue that Hocking and Boser describe identical classifiers or formulas, or even that formulas could be thoughtlessly plucked from each reference and literally combined with mathematical precision to yield the claimed invention. Instead, the Petition explains what aspects of each reference a [skilled artisan] would select and why it would have been obvious for a [skilled artisan] to combine them.

Id. at 16.

4. Patent Owner's Sur-Reply

Patent Owner asserts that Petitioner "attempt[s] to characterize a black box as merely a modular element from which any necessary information can be gleaned" "should be rejected" because "Kohavi itself explains that 'no knowledge of the algorithm is needed, just the interface.'" Sur-reply 5 (citing Ex. 1007, 284). Patent Owner asserts that the SVM weights "are part of the induction algorithm" and undisputedly "*needed* by the feature selection search module in the proposed Kohavi-Boser-Hocking combination," which is directly contrary to Kohavi's principle of operation that no knowledge of the algorithm is needed. *Id.*

Furthermore, Patent Owner asserts that Petitioner's contention that "Hocking's ranking function" would substitute for Kohavi's "feature evaluation" cannot be squared with the wrapper method Kohavi discloses. *See* Sur-reply 6–8. This substitution would completely change the operation of wrapper method taught by Kohavi, Patent Owner argues, because the alleged "Hocking ranking function" does not evaluate the performance of the induction algorithm at all; it simply estimates the "best" subset of the feature set without consideration of how the induction algorithm performed with the feature set. *Id.* at 8 (citing Ex. 1007, 274, 318; Ex. 1009, 531 (Abstract); Ex. 2013, 26:8–14, 29:9–31:6). As such, Patent Owner asserts that "it is unclear how Kohavi's [modified] wrapper method would (or even could) determine which feature set has the 'highest evaluation' as the 'final set on which to run the induction algorithm.'" *Id.* (citing Ex. 1007, 274).

Patent Owner next denies that its arguments improperly depend on bodily incorporation and asserts that they instead account for what the references teach in their entirety including those portions that would lead away from the present invention. Sur-reply 11. Patent Owner criticizes Petitioner for trying to "abstract the teachings of Hocking into a 'ranking principle'" in an effort "to lead the Board to believe that the linear assumptions underlying Hocking's teachings are no longer relevant," which Patent Owner asserts is "far from the truth." *Id.* Patent Owner asserts that since the alleged "ranking principles" are still drawn from mathematical formulas that are based on underlying "linear model assumptions," "so too must the 'ranking principle' obey the 'linear model assumptions.'" *Id.* at 11–12. And because there is no "dispute Boser's SVM does **not** operate under the linear model assumptions of Hocking," Patent Owner reaffirms its argument that a skilled artisan "would **not** have applied the ranking principle of Hocking to Boser's SVM weights as the two methods operate under different model assumptions." *Id.* at 12 (citing Ex. 2013, 71:4–7; Ex. 2014 ¶¶ 35–37).

Patent Owner further criticizes Petitioner's suggestion that a skilled artisan would have known to take the mathematical formulas Hocking uses to justify the "ranking principles" taught and simply "disregard" the error and error variance variables because they address concepts of linear regression that are not applicable to SVMs. Sur-reply 12–13. Patent Owner states, "[i]n this field, persons of ordinary skill in the art do not simply 'disregard' such terms" without justification. *Id.* at 13. Patent Owner asserts that its expert has demonstrated no justification for disregarding the error variance variable exists because a skilled artisan in the field of SVMs would have recognized that an error variance variable is not a constant and still affects the computational result. *Id.* (citing Ex. 2014 ¶¶ 46–54).

Regarding the differing objectives of Hocking and Boser, Patent Owner asserts that, in the Reply, Petitioner "dodges" the pertinent question about whether "Hocking's feature elimination [would] *reduce* the margin between correctly classified training patterns and the SVM's decision boundary that Boser's SVM seeks to *maximize*." Sur-reply 15 (citing Ex. 1008, 1:27-41). Although Petitioner states that the reference's objectives, while admittedly different, are "complementary," Patent Owner asserts that the reader is "left guessing" because Petitioner never explains how the two objectives are complementary. *Id.* (citing Reply 15).

5. *Evaluation of Petitioner's Evidence and Arguments to Show Kohavi, Boser, and Hocking Disclose Ranking the Features According to Their Corresponding Weight Values*

After studying the submissions by both parties, we find Petitioner has not shown persuasively that a skilled artisan at the time of the invention would have had a motivation to combine Kohavi, Boser, and Hocking in the way independent claims 1, 12, and 16 recite. In particular, we are not persuaded by Petitioner's evidence and contention that a skilled artisan would have had a motivation to modify Kohavi's wrapper method to rank the SVM features according to their corresponding weight values as

Elements 1(c), 12(c), and 16(c) recite. For the following reasons, we find Petitioner's evidence and reasoning demonstrates nothing more than a skilled artisan, once presented with the separate pieces of highlighted information in Kohavi, Boser, and Hocking, may have understood that they *could be* combined in the manner claimed. The evidence and reasoning do not demonstrate persuasively that a skilled artisan would have had a motivation to modify Kohavi's wrapper method by *changing the ranking* used in the feature subset selection algorithm *from* an estimation of the performance of an induction algorithm to classify data properly *to* a variable—feature weight—used in the algorithm of an SVM to classify data.

Kohavi establishes that at the time of the invention skilled artisans recognized that "[i]n supervised machine learning, an induction algorithm is typically presented with a set of training instances, where each instance is described by a vector of feature (or attribute) values and a class label." Ex. 1007, 274. It was known that "[t]he task of the induction algorithm, or the inducer, is to induce a classifier that will be useful in classifying future cases" by "mapping from the space of feature values to the set of class values." *Id.* A skilled artisan knew a problem facing induction algorithms was the selection of a subset of features that on which its learning algorithm can focus its attention, while ignoring the rest. *Id.* Kohavi shows a wrapper method was a known method for addressing this problem. *Id.*

According to Kohavi, selecting the appropriate subset for an induction algorithm was understood generally to involve the following steps: (1) "[t]he induction algorithm is run on the dataset" (and the data set is "usually partitioned into internal training and holdout sets, with different sets of features removed from the data"); (2) "[t]he feature subset with the highest evaluation is chosen as the final set on which to run the induction algorithm;" and the resulting classifier (i.e., the induction algorithm) "is then evaluated on an independent test set that was *not* used during the search." Ex. 1007, 274. Kohavi states that "In the wrapper approach . . . the feature subset selection algorithm exists as a wrapper *around* the induction algorithm" and it "conducts a search for a good subset using the induction algorithm itself as part of the function evaluating feature subsets." *Id.* From Kohavi, we find a skilled artisan understood that "[t]he idea behind the wrapper approach . . . is simple: the induction algorithm is considered as a black box." *Id.* (emphasis added); *see also id.* at 284 ("In the wrapper approach . . . the feature subset selection is done using the induction algorithm as a black box (i.e, no knowledge of the algorithm is needed, just the interface)"). As such, Kohavi suggests that the process used by an induction algorithm is irrelevant to the operation of the wrapper approach.

Kohavi's wrapper approach includes a "subset selection algorithm" that "conducts a search for a good subset using the induction algorithm itself as part of the evaluation function." Ex. 1007, 284. The induction algorithm is used with Kohavi's wrapper to provide "induced classifiers," which are evaluated for accuracy "using accuracy estimation techniques." *Id.* Kohavi explains that "[t]he goal of the search is to find the state with the highest evaluation, using a heuristic function" as a guide and, because the actual accuracy of the induced classifier is not known, an accuracy estimation is used as the evaluation function. *Id.*

Kohavi demonstrates that at the time of the invention a skilled artisan was aware that the wrapper approach may be implemented with an induction algorithm to search for a good feature subset. Ex. 1007, 274, 284. Kohavi discloses a wrapper method that includes a process for, among other things, inducing the induction algorithm to provide a classifier to be evaluated in furtherance of the search. *Id.*; Ex. 1003 ¶ 199; Ex. 2014 ¶¶ 10–18. As part of this evaluation, Kohavi teaches a wrapper method that evaluates the accuracy of classifiers provided by an induction algorithm and, based on the classification performance of the induction algorithm, decides whether to keep or discard the feature subset. Ex. 1007, 274, 284; Ex. 1003 ¶ 199; Ex. 2014 ¶¶ 10–18. A skilled artisan would have understood that this evaluation process in Kohavi's wrapper method includes computing a ranking criteria using the classification performance of the induction algorithm to properly classify a feature subset. *Id.*

Kohavi does not disclose or suggest, however, that a wrapper method uses the feature weights generated by an SVM as a ranking criteria or that the SVM feature weights may be used to compute a ranking criteria for selecting a feature subset in the wrapper approach. Importantly, Petitioner does not argue otherwise. *See generally* Pet.

We agree with Petitioner that a skilled artisan at the time of the invention would have known that an SVM, such as the one disclosed by Boser, is a kind of "learning machine" designed to "to learn to recognize as well as possible an unknown pattern, given a set of training examples." Pet. 28 (citing Ex. 1008, 1:9, 1:12-19; Ex. 1003, ¶211). A skilled artisan at the time of the invention would have known that Boser's SVM was a kind of induction algorithm that receives a set of training instances, where each instance is described by a vector of feature (or attribute) values and a class label, to formulate a decision function from the training set. *Id.* SVMs were known by skilled artisans to experience the computational burden problem of having to select of a subset of features that its learning algorithm can focus its attention, while ignoring the rest, in order to more effectively formulate a decision function and avoid the poor generalization and overfitting issue that may occur with the inclusion of irrelevant features. *Id.* Because Kohavi teaches that its wrapper method may be used successfully to address the computation burdens associated with selecting an optimal subset of features enabling the learning machine/induction algorithm to create a more accurate classifier, we are persuaded that a skilled artisan would have known to combine the Kohavi wrapper method with an SVM, such as the one Boser discloses.

Even in combination, however, the teachings of Kohavi and Boser still fail to disclose or suggest ranking the features according to their corresponding weight values. Petitioner notably does not contend otherwise. *See generally* Pet. We find that the resulting combination of Kohavi and Boser, after viewing the teachings as a whole, only supports a determination that it would have been obvious for a skilled artisan to modify Kohavi's wrapper method to specifically include an SVM as the induction algorithm for which the feature subset selection algorithm wraps around. As such, even when combined, the disclosed wrapper method continues to compute ranking criteria using the performance of the induction algorithm, which is treated as a "black box" (i.e, no knowledge of the algorithm is needed, just the interface), to classify a feature subset; only, in this case, the ranking criteria is of the performance of the induced classifier produced by Boser's SVM.

The issue therefore becomes whether the addition of Hocking's teachings demonstrates that a skilled artisan would have had a motivation at the time of the invention to modify Kohavi's wrapper method, when used with a Boser SVM, to change the ranking *from* the estimated accuracy or performance of Boser's SVM classifier *to* the weights computed by Boser's SVM algorithm in the process of providing a classifier. As noted above, at most, the combined Kohavi/Boser/Hocking disclosures suggest that a skilled artisan, once presented with the separate pieces of highlighted information from those references, *may* have understood that they *could be* combined in the manner claimed, but that is not enough because Petitioner has not shown persuasively why a skilled artisan *would have known* to pick out those three references and combine them to arrive at the claimed invention. *See Personal Web Techs., LLC v. Apple, Inc.*, 848 F.3d 987, 993 (Fed. Cir. 2017) (holding that the Board's reasoning that "a person of ordinary skill in the art reading Woodhill and Stefik would have understood that the combination of Woodhill and Stefik would have allowed for the selective access features of Stefik to be used with Woodhill's content-dependent identifiers feature" to be deficient to establish a motivation because it "say[s] no more than that a skilled artisan, once presented with the two references, would have understood that they could be combined. And that is not enough: it does not imply a motivation to pick out those two references and combine them to arrive at the claimed invention").

Hocking "is directed to the problem of selecting the best subset of independent variables in a multiple linear regression problem." Pet. 26 (citing Ex. 1009, 531). Hocking addresses this problem by computing ranking criteria based on estimates of a weight vector and iteratively removing one or more variables with the minimum ranking criteria until the subset providing the "best regression" is identified. *Id.* at 27 (citing Ex.

1009, 533; Ex. 1003, ¶ 207). Hocking computes that ranking criteria based on a classical "linear model" using "well known" linear model assumptions to allow "estimates with many optimality properties [to be] obtained by minimizing the residual sum of squares." Ex. 1009, 531.

For Elements 1(c), 12(c), and 16(c), Petitioner contends that "[i]n the context of the combination [of Kohavi/Boser/Hocking], a feature subset is evaluated using Boser's SVM as the induction algorithm, and ranking criteria [is] computed based on the estimates of feature weight values determined *according to the teachings of Hocking*." Pet. 48, 57–58 (emphasis added). This, however, only begs the central question of *why* a skilled artisan had a motivation to change the ranking criteria in Kohavi's wrapper method *from* one that is based on estimates of how well the induction algorithm (an SVM in the proposed combination) classifies data (i.e., an evaluation of the SVM result) *to* one that is based on the feature weight estimates computed by an SVM to classify data (i.e., an evaluation of a variable used to provide an SVM result).

Petitioner states that a skilled artisan "would have been motivated to use Hocking's simple and computationally efficient ranking based on weight values" generated by Boser's SVM, but Petitioner provides no persuasive basis to find that a skilled artisan would have viewed Hocking's "ranking based on weight values" to be a simple or computationally efficient ranking criteria in Kohavi's wrapper method used with an SVM. Although Petitioner's expert refers to Hocking's teaching that an optimal subset of variables can be derived "with a minimum of computation" (Ex. 1003 ¶ 216), we find little persuasive value in this representation as a motivating factor because it is unclear a skilled artisan would have recognized a link between the minimal computational benefits of Hocking's ranking when applied to Kohavi/Boser combination. It is undisputed that Hocking's ranking is characterized as using "a minimum of computation" in the context of a very different application than SVMs and relies on different objectives and assumptions. *See* Reply 11, 13, 15. Petitioner has not shown persuasively that those computational gains would also have been recognized by skilled artisans to occur in the context of Kohavi's wrapper method applied to Boser's SVM.

Petitioner expert *suggests* that, because Kohavi describes using the computationally complex "backwards feature elimination," "Hocking's teachings of simple and computationally efficient ranking criteria for feature selection" would have motivated further improvement to Kohavi's wrapper method." Ex. 1003 ¶ 216. No persuasive technical reasoning or evidence, however, is offered to actually demonstrate that a skilled artisan would have reasonably expected an improvement to occur; for example, neither Petitioner, nor its expert, explains *how* the modification would have improved upon what was previously known. All that is provided is a conclusory statement that has little, if any, evidentiary value.

Additionally, Patent Owner makes two points that diminish the persuasiveness of Petitioner's contention that Hocking's ranking would have been recognized to be simple or computationally efficient in the context of the Kohavi's wrapper method and Boser's SVM. First, although cast as a principle of operation issue, Patent Owner highlights Kohavi teaches that its wrapper approach "exists as a wrapper *around* the induction algorithm" with the idea being the induction algorithm is a black box (i.e, no knowledge of the algorithm is needed, just the interface). Ex. 1007, 274, 284 (emphasis added). Setting aside whether Petitioner's proposed modification changes Kohavi's *principle* of operation, the proposed modification would indeed change the operation (principle or not) of Kohavi's wrapper method to one that is dependent upon knowledge of the algorithm that an induction algorithm uses, because the wrapper method would have to have knowledge of what variables are computed in the process of providing a classifier. We are not persuaded that adapting a wrapper method to the specific algorithm of an induction algorithm in order to rank the features according to their corresponding weight values would have been viewed by skilled artisans to be simpler than (or even as simple as) rank the features according to the results provided by an induction algorithm and evaluating its performance.

Regarding Petitioner's contention that such modification would have been more computationally efficient, Patent Owner's expert credibly and persuasively explains why a skilled artisan *would not have recognized* Hocking's linear regression model to be reducible to the computationally efficient form that Petitioner's expert suggests. Ex. 2014 ¶¶ 40–54. In particular, we find it credible that a skilled artisan would not have known at the time of the invention that the estimated variance values of the combined Hocking's Equations (5) and (6) could be treated as a constant and, as a result, canceled out of the equations. *Id.* In view of this fact, we are not persuaded that Hocking's reference to simplicity and computational efficiency demonstrates a skilled artisan had a motivation to make the modification Petitioner proposes. Furthermore, because a skilled artisan would not have recognized at the time of the invention that Hocking's Equations (5) and (6) were reducible to the form Petitioner proposes, Petitioner's contention that the proposed modification would have yielded predicable results is also unpersuasive.

Petitioner notes that "Kohavi uses accuracy as ranking criteria, [but] explains that "one can trivially use a cost function instead of accuracy as the evaluation function for the wrapper." Pet. 31 (citing Ex. 1007, 309). Petitioner summarily concludes that "using Boser's feature weight values as the evaluation function of Kohavi's wrapper method, according to the teachings of Hocking, would [likewise] be a trivial substitution of known methods." *Id.* (citing Ex. 1003 ¶ 215). However, neither Petitioner nor its expert provide any evidence or technical reasoning that demonstrates a skilled artisan would have connected the "cost function" referenced by Kohavi to the use of Boser's feature weight values for the evaluation function. In fact, Petitioner's expert declaration simply parrots the Petition in this respect. *Compare* Pet. 31 *to* Ex. 1003 ¶ 215. And there is no persuasive support at page 309 of Kohavi for Petitioner's reasoning; to the contrary, page 309 of Kohavi suggests that the referenced "cost function" relates to "misclassification costs," which is just another way of evaluating the classification *result* produced by an induction algorithm. We are not persuaded Kohavi suggests more generally that any cost function associated with an induction algorithm would be a "trivial" modification.

There is also little persuasive weight afforded to Petitioner's contention that a skilled artisan would have been motivated to modify Kohavi's ranking criteria to estimated feature weights because Kohavi, Boser, and Hocking all involve techniques that are based on similar classical statistic principles. *See* Pet. 32–33. This contention "say[s] no more than that a skilled artisan, once presented with the [three] references, would have understood that they could be combined. And that is not enough: it does not imply a motivation to pick out those two references and combine them to arrive at the claimed invention" *Personal Web Techs.*, 848 F.3d at 993. Furthermore, in applying the classical statistical principle Hocking teaches to this case, we agree with Patent Owner that Petitioner relies on erroneous assumptions, which a skilled artisan would not have made, that undermines the persuasiveness of Petitioner's contention. And if the classical statistic principles Hocking does disclose were applied, it is not clear that the feature subset selected for use with an SVM would provide accurate classifications. *See* Ex. 2014 ¶¶ 51–55.

When all of Petitioner's evidence and arguments regarding the combination of Kohavi, Boser, and Hocking are viewed in light of Patent Owner's arguments, we are not persuaded that using the weights generated by an SVM to rank the features according to their corresponding weight values in Kohavi's wrapper method would have been a known (or suggested) technique to skilled artisans at the time of the invention for selecting a subset of features. Hocking teaches the use of weights to select features in the context of a different application than the one presented by Kohavi/Boser combination, moreover, Hocking's teaching is based on objectives and assumptions that are inapplicable to the Kohavi/Boser combination. Petitioner has not shown persuasively that a skilled artisan would have recognized Hocking's technique to be readily applicable to both applications. Petitioner has not provided sufficient evidence or technical reasoning to demonstrate persuasively that the weight ranking criteria used in Hocking would have been a known technique for selecting a feature subset in Kohavi/Boser combination. Petitioner provides no evidence that a skilled artisan would have recognized that the weights generated by an SVM to classify a feature subset may also be evaluated to rank feature subsets to optimize the operation of an SVM. In fact, Petitioner fails to even establish more generally that it was known that an SVM's feature weights may be used as ranking criteria to select the feature subsets to be run with an SVM. There is therefore no persuasive basis to find a skilled artisan knew at the time of the invention that an SVM's feature weights may be used to address the computation burdens associated with selecting an optimal subset of features enabling the learning machine/induction algorithm to create a more accurate classifier.

As a result, we find that Petitioner has not persuasively shown that Kohavi, Boser, and Hocking disclose Elements 1(c), 12(c), and 16(c) in challenged independent claims 1, 12, and 16. Therefore, after reviewing all of the evidence of record, including Petitioner's cited evidence and the declarations of Dr. Evgeniou, we determine that Petitioner has not shown by a preponderance of the evidence that independent claims 1, 12, and 16 are unpatentable for being obvious in view of Kohavi, Boser, and Hocking. Because challenged claims 5, 6, and 10, respectively, depend from claim 1, we therefore likewise determine that Petitioner has not shown by a preponderance of the evidence that claims 5, 6, and 10 are unpatentable.

F. Additional Obviousness Grounds

Petitioner asserts an additional obviousness ground based on Kohavi, Boser, Hocking, and Cristianini, challenging claims 1, 5, 6, 10, 12, and 16. Pet. 64–69. This additional ground still relies on Petitioner's proposed combination of Kohavi, Boser, and Hocking to show Elements 1(c), 12(c), and 16(c), and does not remedy the deficiency explained above. *See id.*

Therefore, Petitioner has not met its burden to show that claims 1, 5, 6, 10, 12, and 16 would have been obvious over the asserted prior art.

III. CONCLUSION

We have reviewed the Petition, Patent Owner Response, Petitioner Reply, and Patent Owner Sur-Reply. We have considered all of the evidence and arguments presented by Petitioner and Patent Owner, and have weighed and assessed the entirety of the evidence as a whole. For the reasons above, we determine, on this record, that Petitioner has not demonstrated by a preponderance of evidence that claims 1, 5, 6, 10, 12, and 16 of the '959 patent are unpatentable over the asserted prior art.

Claim(s)	**35 U.S.C. §**	**Reference(s)/Basis**	**Claims Shown Unpatentable**	**Claims Not Shown Unpatentable**
1, 5, 6, 10, 12, 16	103	Kohavi, Boser, Hocking		1, 5, 6, 10, 12, 16
1, 5, 6, 10, 12, 16	103	Kohavi, Boser, Hocking, Cristianini		1, 5, 6, 10, 12, 16
Overall Outcome				1, 5, 6, 10, 12, 16

ORDER

It is hereby: ORDERED that Petitioner has not shown that claims 1, 5, 6, 10, 12, and 16 of the '959 patent are unpatentable under 35 U.S.C. § 103(a) as obvious over the asserted prior art; and

FURTHER ORDERED that this Decision is final, and a party to this proceeding seeking judicial review of the Decision must comply with the notice and service requirements of 37 C.F.R. § 90.2.

For PETITIONER:

Lori Gordon
Lauren May Eaton
PERKINS COIE LLP
gordon-ptab@perkinscoie.com
eaton-ptab@perkinscoie.com

For PATENT OWNER:

Tarek Fahmi
Jonathan Tsao
ASCENDA LAW GROUP, PC
tarek.fahmi@ascendalaw.com
jonathan.tsao@ascendalaw.com

Eleanor Musick
MUSICK DAVISON LLP
eleanor@mdiplaw.net

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

INTEL CORPORATION,
Petitioner,

v.

HEALTH DISCOVERY CORPORATION,
Patent Owner.

IPR2021-00552
Patent 7,542,959 B2

Before LYNNE H. BROWNE, GARTH D. BAER, and FREDERICK C. LANEY, *Administrative Patent Judges*.

BAER, *Administrative Patent Judge*, dissenting.

I respectfully dissent from the majority's decision finding none of the challenged claims are unpatentable.

The majority finds Petitioner did not show a sufficient motivation to implement Hocking's weight-vector ranking criteria in Kohavi's wrapper-based method. *Supra* Section II.E.5. I disagree. Petitioner explained, with support from its expert, Dr. Evgeniou, that its proposed addition of Hocking's vector weight ranking criteria "applies a known technique (Hocking's variable selection) to a known device (Kohavi's RFE method using Boser's SVM) which is ready for improvement to yield predictable results." Pet. 32 (citing *KSR*, 550 U.S. at 417; Ex. 1003 ¶ 217). As the Supreme Court has explained, "[t]he combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results." *KSR*, 550 U.S. at 416. I agree with Petitioner that the claimed invention is an obvious combination of known techniques applied to a known device, yielding only predictable results and thus obvious under *KSR*'s framework.

In my opinion, Petitioner has shown sufficiently that the challenged claims of the '959 patent are unpatentable.